Essays on Food Demand and Food Retail Competition in Local Markets

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Yanghao Wang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTORATE OF PHILOSOPHY

Metin Çakır

July, 2018

Acknowledgment

I would like to thank my advisor and dear friend, Metin Çakır. Through his thoughtful guidance, unwavering support, and sincere encouragement, Metin has been a major positive influence throughout my Ph.D. studies. I deeply value our discussions, whether they involve research or simply our life stories. I would not complete my dissertation in two years without working with him.

I would like to express my gratitude to the members of my advisory committee—Michael Boland, Hikaru Peterson, and Chengyan Yue—for their valuable input. I was fortunate to have worked in an office next to Michael, whose humorous words and encouragement made my days as a graduate student much more pleasant.

I am grateful to Marc Bellemare, Terry Roe, and Rodney Smith for their generous advice. I am also thankful to the Department of Applied Economics at the University of Minnesota for financial assistance.

I genuinely appreciate the fellowship with Yuan Chai, Yinxiao Chu, Andrew Keller, and Shuoli Zhao. Their support has helped me advance in my research.

Finally, I am grateful to my parents and parents-in-law for supporting my desire to pursue a Ph.D. I am eternally indebted to my considerate wife and chief cheerleader, Hangyang Xue, for her optimism and unyielding support throughout the past four years.

Dedication

To my grandfather

Qixiang Wang

Abstract

Consumers become more health-conscious and have higher and more diverse expectations for food quality, ranging from food taste and nutrition features to the characteristics of food production, processing, and marketing. In response, large numbers of new products are introduced and a wide range of changes have occurred in food and agricultural markets. Accordingly, this dissertation comprises three essays investigating questions relevant to these changes and providing implications for food policy and retailing.

The first essay focuses on the increasing popularity of the "New Super Grain"—Ethiopian teff. Specifically, we examine Ethiopian consumers' welfare losses due to increasing teff prices and evaluate the effectiveness of alternative food aid policies in alleviating these losses. Using data from two waves of Ethiopia Socioeconomic Survey 2013-2014 and 2015-2016, we estimate a two-stage demand system and document the consumption patterns of cereals in Ethiopia. We find that teff is the most own-price inelastic grain in the cereal market and a one percent increase in teff prices leads to a decrease of 0.38 percent in total consumer welfare. Subsequently, our results suggest wheat aid is an effective policy in reducing the impacts of increasing teff prices, which lends support to the ongoing Ethiopian policy that distributes subsidized wheat on a large scale.

The second essay focuses on the introduction of new demand-enhancing agricultural products. Specifically, we evaluate the welfare impacts of the introduction of Honeycrisp apples. We estimate structural models of consumer demand and retailer competition using store scanner data covering 61 cities across the United States during the period from March 2009 to February 2015. The results show that, on average, the introduction of Honeycrisp apples increases consumer welfare by 3.14 cents per pound, of which 2.98 cents is explained by the increased number of total apple varieties and 0.16 cents by the decline in prices of competing apple varieties. The results also show that the introduction of Honeycrisp apples has increased the total sales quantity by 8.03 percent and the total sales revenue by 21.25 percent over the study period.

The third essay examines the food retail competition in local markets by addressing the heterogeneity in households' choice sets of stores, shopping baskets, and travel distances. A revised mixed logit model is developed to model the household choice of shopping stores that enable us to calculate stores' price elasticities and recover their gross profit margins under alternative pricing strategies. We construct a dataset for estimation by matching the information in 2016 IRI household and retail scanner datasets. The results show that without considering household travel distance for shopping, we might overestimate stores' price elasticities and underestimate their gross profit margins. The results also suggest that households prefer to visiting closer stores at expense of paying higher prices for their shopping baskets. Finally, we find that one increase in the number of nearby rivals within 5 kilometers from a store is associated with a decrease of 1.6 to 2.4 percent in the store's gross profit margin depending on different pricing strategies.

Contents

List of Tables

List of Figures

Chapter 1. Introduction

The effective functioning and organization of food and agricultural markets are vital to improve people's lives and, consequently have been central to research and policy goals throughout history. Over the past century, a tremendous effort has been put to the development of agricultural technologies and processes that enable us to increase production and alleviate the risk of hunger and malnutrition to a large extent. Although such problems still exist, our attention has shifted away from the challenges in food and agricultural production and moved towards issues pertinent to consumption (Lusk, Roosen, and Shogren 2011). As consumers have increasingly high and diverse expectations for food quality, food and agricultural markets are becoming more consumer-oriented (Unnevehr et al. 2010).

In the last decades, a wide range of changes has taken place. Food and agricultural products have been gradually more differentiated to serve consumers' heterogeneous tastes and expectations. Consumers demand for quality ranges from taste, appearance, and healthfulness of food to how food is produced, packaged and marketed. In response, an abundance of new food products is introduced, providing consumers more options than ever. These changes have reshaped the landscape of food and agricultural markets and accordingly raised public interests in timely empirical research on consumer demand for foods and implications for food policy and retailing.

This dissertation consists of three essays focusing on research questions in consumer demand for foods and consumer choices of stores. Overall, each essay builds an analytical framework centered on consumer demand with a strong connection to economic theory. The first essay investigates the increasing popularity of teff consumption in developed countries on consumer welfare in Ethiopia. Headlined by the popular press, there is an increasing demand for ancient grains in Europe and North America. Teff is one of such grains and is primarily grown in Ethiopia. With two to three times as much iron and calcium as quinoa—"Super Grain"—contains, teff is entitled the "New Super Grain." Having concerns about a dramatic increase in domestic teff prices driven by the global demand, the Ethiopian government imposed an export ban on teff grain and flour in 2006. This trade restriction was expected to stabilize teff prices and enhance food security in Ethiopia. After about a decade, in 2015, the government partially lifted the export ban and, consequently, teff prices increased. The average teff price in 2016 and 2017 was 10 percent higher than the price in 2014 and 2015. In the meantime, the government had been implementing a wheat import program since 2008, which has enabled consumers to substitute teff with wheat. The first essay measures the loss of consumer welfare due to the increasing teff prices and the impacts of food aid policies in mitigating this loss. The results provide new evidence on the welfare impacts of price increases and improve our understanding of the cereal market and food policy in Ethiopia.

The second essay focuses on the introduction of new demand-enhancing agricultural products. The agriculture sector in the United States (U.S.) has introduced to the market more than 3,500 new varieties of fruit and vegetables since 2011. This large-scale introduction is primarily driven by research programs toward consumer-oriented products that are financially supported by agricultural research and development (R&D) investments. At the same time, the development of patent protection laws gave universities the permission to attain the ownership of inventions made with federal funding and, thereby, enabled them to finance agricultural research by transferring a part of their patent rights to the private sector. Therefore, it is of interest to know the economic benefits of new demand-enhancing agricultural products. The second essay examines welfare impacts from the introduction of a new agricultural product, with a focus on the case of Honeycrisp apples. Honeycrisp apples are one example of the most successful products developed by the public university system. The results quantify the economic benefit of Honeycrisp apples for both consumers and retailers and provide some implications for agricultural R&D initiatives in the public university system.

The third essay looks into store competition in local food markets and the determinants in the household choice of stores. According to the 2009 National Household Travel Survey, the annual average of a personal trip for shopping was 6.5 miles, suggesting that consumers might be reluctant to travel a long distance for their shopping. Ideally, they would like to shop in a store, both close to their residential communities and selling the products at lower prices. As a result, stores probably compete against nearby rivals by providing differentiated products in local markets. Through efforts such as developing and equipping stores, the U.S. Healthy Food Financing Initiatives expect to increase food access in underserved communities. However, an increasing number of stores do not necessarily imply an improvement of food access, since household decisions on food consumptions are jointly determined by the accessibility and affordability of stores. On one hand, opening a new store will directly increase the accessibility of food and raise competition in the local market and in turn reduce food prices in the local market. On the other hand, stores in the local market might cooperate with their nearby rivals and thereafter exercise market power and maintain considerable profit margins. Therefore, the third essay examines the store competition in local markets by addressing the heterogeneity in household choice sets of stores, travel distances, and shopping baskets. The results help us understand household valuations of store proximity and the extent to which retailer concentration drives up local food prices. The results also contribute to the discussion on community economic development policy and anti-trust regulation in local markets.

Chapter 2. The Welfare Impacts of Increased Demand for The "New Super Grain"

2.1. Introduction

In the last decade, there is an increasing trend in demand for ancient grains such as quinoa and teff in Europe and North America. These grains have been preserved in the natural states for thousands of years but recently became popular due to their nutritional and environmental value.[1] The recognition from international organizations also fuels this growing popularity of ancient grains. For example, the United Nations launched the "International Year of Quinoa" in 2013 to raise the awareness of the grain, as well as its value in culture, economics, and health. With two to three times as much iron and calcium as quinoa contains, teff is entitled the "New Super Grain" (Jeffrey 2015; Secorun 2016).

Teff, one of the main staples in Ethiopia, provides more than 15 percent of all calories consumed in the country (O'Connor 2016). The increasing global demand for teff raised the public concerns about the affordability of teff and food security in Ethiopia as well. In response, the Ethiopian government imposed an export ban on teff grain and flour in 2006 to stabilize the domestic teff price (Provost and Jobson 2014; Jeffrey 2015; Nurse 2015; O'Connor 2016; Secorun 2016).[2] However, the Food and Agriculture Organization of the United Nations (FAO) challenged the purpose of this restriction, arguing that the poor agrarian households in rural Ethiopia primarily grow teff as a cash crop and, thus, would benefit from higher teff prices. The FAO reports suggest that wheat, maize, and sorghum are more critical than teff to enhance food security (Demeke and Di Marcantonio 2013a; Assefa 2015).

---

[1] The terms "grain" and "cereal" will be used interchangeably through this chapter.

[2] A wide array of policies designed to control food prices are frequently employed in developing countries, since food expenditure in these countries accounts for a large share of household consumption (Porteous 2017). In the poorest countries, household diets heavily depend on staple grains rather than meat or processed foods (Abbott 2012).

Due to improvements of technology and government policies, the teff production in Ethiopia doubled between 2006 and 2015. The growth of production helped to alleviate concerns about maintaining a sufficient teff supply in the domestic market. As a result, the government began to lift the teff ban partially by granting export licenses to 48 commercial farmers (O'Connor 2016; Secorun 2016).[3] The export permission was put into effect in July 2015, the beginning of the fiscal in Ethiopia. After the withdrawal of the ban, teff prices in Ethiopia increased considerably. The average price of teff in 2016 and 2017, on average, was 10 percent higher than its level in 2014 and 2015 (Ethiopian Grain Trade Enterprise 2017). Since teff is a primary staple in Ethiopia, the increase in its prices would inevitably hurt consumer welfare and have adverse impacts on food security. Meanwhile, the Ethiopian government maintained a large quantity of wheat import since 2008 and distributed the imported wheat at subsidized prices in the market (Demeke and Di Marcantonio 2013b). Consumers are hence able to substitute teff with wheat. However, the overall change in consumer welfare is unclear.

To improve our understanding of the cereal market in Ethiopia, this chapter investigates the following questions. How large is the adverse impact of the increasing teff prices on the consumer welfare of an average Ethiopian household? How effective is the ongoing food aid of wheat in reducing this adverse impact? Would it be more effective if less expensive cereals, such as maize, barley, and sorghum, are provided as alternative food aids? To answer these questions, we particularly examine the food expenditure in Ethiopia and the associated cereal consumption patterns. A two-stage food demand system is estimated using the aggregated consumption data from the two waves of Ethiopia Socioeconomic Survey 2013-2014 and 2015-2016. With the estimated demand system, we are able to calculate price and expenditure elasticities of grains,

[3] According to the Ethiopian Agricultural Transformation Agency, the government would expand the applicability to cooperatives of smallholder farmers, if this pilot program was successful (O'Connor 2016; Secorun 2016).

evaluate the change in consumer welfare due to the increasing teff prices, and explore the role of alternative food aid policies in reducing the impact of "teff fever."

The cereal production is an essential component of Ethiopian agriculture and has been continuously subject to public policies over the past fifty years (Rashid 2010). The cereal sub-sector employs 60 percent of rural labor and provides more than 60 percent of household calorie consumption with an equivalent value at more than 40 percent of household food expenditure (Rashid 2010). Teff is the most valuable staple grain in Ethiopia. It is grown on 23 percent of the overall cultivated land by 43 percent of agrarian households who raise grain crops (Central Statistical Agency 2016). It is the only grain in which Ethiopia has a comparative advantage in the international cereal market (Demeke and Di Marcantonio 2013a).

Most of the literature on the impacts of price change in the cereal market examines the adverse impact of food price spikes on the well-being of households in poor countries. Ivanic and Martin (2008) investigate the short-run effects of higher staple food prices on poverty in low-income countries. According to household food consumption and production data from nine low-income countries as well as international market prices, they find the increasing food prices generally raises the overall poverty, although the impacts might differ by commodity and by country. On the contrary, using data from the household survey and the information of domestic prices in Vietnam, Vu and Glewwe (2011) find average welfare gain for net sellers outweighs average welfare loss for net buyers and hence conclude that on average higher food prices contribute to the increase of household welfare in Vietnam.

The only two economic studies on the effects of rising prices of "super grains" are Stevens (2015) and Bellemare et al. (2016). Both of them investigate the problems triggered by the "quinoa fever" since 2009. By modeling household demand for foods, Stevens (2015) finds the rising quinoa prices do not affect household nutrition outcomes through the pathway of regional preferences for quinoa. The author attributes this insignificance to the small share of quinoa in local household diets. Bellemare et al. (2016) examine the welfare impacts of rising quinoa prices in Peru, using pseudo-panel and the difference-in-difference method. They find a positive association between the price of quinoa and household welfare within the areas where quinoa is consumed, and a robust positive relationship between household welfare and household quinoa production as well. In contrast to the small share of quinoa in household diets in South America, teff is a local staple food in Ethiopia, accounting for a large proportion of household food consumption, especially for urban residents.

This chapter is closely related to the literature on cereal supply and demand in Ethiopia. Mottaleb and Rahut (2018) investigate the determinants of household production and consumption of teff using the first two waves of the Ethiopia Socioeconomic Survey. Their results imply a strong value chain of teff between farmers and consumers, suggesting to link producers with international markets. Their study, however, does not evaluate the welfare impact of the increasing teff prices on consumer welfare. In another relevant study, Tefera et al. (2012) examine the food demand and consumer welfare in Ethiopia. They estimate a household demand system for foods using six waves of the Ethiopia Rural Household Survey from 1994 to 2009. They find the rising food prices benefit both net-cereal sellers and autarkic and net-cereal buyers. There are two distinctive features in this study. First, we use updated survey data that are nationally representative for both rural and urban areas. Second, we model the decision process of food consumption in a two-stage approach and focus our discussion on the consumption patterns within the segment of cereals.

This chapter also enriches the literature on the trade restriction in agricultural markets. Most of the existing studies have found the adverse impact of trade restrictions. Martin and Anderson (2012) estimates the impacts of domestic market-insulating policy on international price spikes and find that this policy intervention makes the problem worse (i.e., not only fail to stabilize prices in the domestic market but also contribute price volatility in the international market). A sizeable literature is in line with this point of view after investigating the impacts of export restrictions (i.e., extra export tax, export quota, and export ban) imposed by grain exporting countries during the world food price crisis in 2007 and 2008. For example, Götz et al. (2013) focus on export restrictions in Russia and Ukraine during the crisis and argue the wheat producers experienced a loss of welfare and the government interventions destabilized the wheat market. Using a 10-year period data from 5 countries in Eastern and Southern Africa, Porteous (2017) investigates the effect of 13 short-term export bans on maize. His results show that the export bans do not have a significant effect on cross-border gaps in maize prices and even destabilize markets by signaling price increases rather than preventing them. In contrast to these studies, we discuss the impact of the withdrawal rather than the imposition of a trade restriction. Moreover, we evaluate the effectiveness of alternative food aid policies in mitigating the consumer welfare loss due to the increasing teff prices and measure the distance of these policies to the ideal one that offers households an equivalent cash transfer.

The rest of this chapter is organized as follows. In Section 2.2, we introduce some background information about teff and the Ethiopian cereal market. Section 2.3 displays the data. Section 2.4 presents the analytical framework and the estimation approach. In Section 2.5, we discuss the results of our model and welfare analysis and then conclude in Section 2.6.

2.2. Teff in Ethiopia

Teff is a gift from Ethiopia to the world, which is gluten-free and high in protein, iron, and calcium. It has been grown and consumed by Ethiopians for millennia and recently become an upper-class staple grain for health-conscious consumers in Europe and North America. Resistant to many biotic and abiotic stresses, teff is highly adaptable to various terrains from basin to highland (Assefa 2015). According to the Central Statistical Agency (CSA) annual reports on the Agricultural Sample Survey in Ethiopia from 2005/2006 to 2015/2016, teff is the second most widely grown crop, accounting for the largest share of farmland over the past decade. Figure 2.1 shows the number of households who grew teff increased from 5.2 million in 2005 to 6.6 million in 2015, and the farmland used for teff rose from 2.2 million hectares in 2005 to 2.8 million hectares in 2015. On average, the farmland used for teff was 30 percent more than that for maize, while the production quantity of teff was 50 percent less than that of maize. Among the primary cereals, teff was the least productive one, though the average yield of teff had been increasing over the last 10 years up to 15.6 quintals per hectare. In 2015, the average yield of teff was less than half as much as that of maize and was about sixty percent of the average yield of wheat and sorghum. The yield of teff is doomed to be lower since teff has been sheltered in Ethiopia for thousands of years and seldom exposed to agricultural research and investment (Jeffrey 2015). For the harvesting purpose, the tiny teff seeds, as small as poppy seeds, are not applicable to modern agricultural machinery designed for other staple crops.

The Agricultural Marketing Corporation (AMC), a government parastatal, controlled the grain trade in Ethiopia until the 1990s. In 1991, because of the policy reform, the AMC was replaced by the Ethiopian Grain Trade Enterprise (EGTE). The EGTE was expected to compete with the private sector in the open market as a representative of the government. However, it does not intervene in the teff market since teff is not a cereal included in food aid in Ethiopia (Assefa 2015). Figure 2.2 shows the wholesale prices for primary cereals from 2014 to 2017 reported by the EGTE. Accordingly, before the export restriction was partially withdrawn in mid-2015, the price of teff was fluctuating between 9 and 10 Birrs per kilogram, more than twice the price of maize and 20 to 50 percent higher than the prices of wheat, barley, and sorghum. After 2015, the price of teff rose up and reached as high as 11.5 Birrs per kilogram in October 2016. By contrast, the price of wheat declined from 7 Birrs in 2014 to about 5 Birrs per kilogram in 2017. Meanwhile, the price of barley varied slightly around 6 Birrs per kilogram, the price of sorghum experienced a big variation ranging from 3.9 Birrs and 6.7 Birrs per kilogram, and the price of maize stayed between 3 Birrs and 4 Birrs per kilogram. It is worth noting the nominal cereal prices are substantially affected by the food price inflation rate and all the wholesale prices in Figure 2.2 are deflated by the CSA consumer price index (CPI) for food and non-alcoholic beverages. Figure 2.3 displays the CPI data and shows that Ethiopia has been undergoing a food price inflation, especially in 2012 at an increasing rate of 15 percent and in 2015 at 12 percent.

2.3. Data

The data are obtained from the Ethiopia Socioeconomic Survey (ESS), a collaborative project conducted by the Central Statistical Agency of Ethiopia (CSA) and the World Bank Living Standards Measurement Study (LSMS) team. The first wave was implemented in 2011-2012, including 3,969 households from 333 enumeration areas (EAs), of which 290 are rural EAs and 43 are small town EAs.[4] An EA is a geographic area delineated for census data collection in Ethiopia, usually including 150 to 200 households. In order to represent the full population in Ethiopia, the second wave in 2013-2014 and the third wave in 2015-2016 added additional 100 middle- and large-sized town EAs and consequently expanded the sample size to 5,469 households. A two-stage sampling method is applied, wherein EAs and households are selected successively in turn. The data include the information of household sampling weights for aggregation. In the end, there were 5,262 and 4,954 households from 433 and 432 EAs in the second and the third wave of ESS, respectively, with an average of 11 to 12 households per EA. These EAs are sampled from 75 geographic zones across 11 regions in Ethiopia except the southeastern part of the country, where there is semi-desert or desert (see Figure A1). To maintain a cohesive and nationally representative sample, we only use the two waves of ESS in 2013-2014 and 2015-2016.

[4] The first wave of ESS can be regarded as an "upgraded" version of the Ethiopia Rural Household Survey (ERHS), a well-known project conducted by multiple research institutes in 1994, 1995, 1997, 1999, 2004, and 2009.

The fieldwork of ESS begins every September. The household- and community-part of ESS are distributed and collected between February and April in the next year. Households report their food consumption quantity and expenditure by item (e.g., teff, wheat, barley, maize, sorghum, horse beans, potato, and etc.) in the past week prior to the survey interview. Food prices are consequently obtained by dividing the associated quantity of purchase into the item-specific expenditure. However, food prices could be missing, if there is no any record of household expenditure on associated items in the past week. To address this challenge, we also use the community-part of ESS, which includes local retail prices by food item that are directly observed from up to two nearby markets.[5]

In this chapter, we aggregate the data at the EA level. In particular, an item-specific price in an EA is an average over associated prices reported by local retailers.[6] An item-specific expenditure is the item-specific price multiplied by the aggregated quantity, which is obtained by adding up household consumption quantities with sampling weights.[7] In turn, the expenditure shares by item are able to be calculated. All prices are deflated by the CSA CPI.

Table 2.1 presents the sample statistics of expenditures, prices, quantities, and expenditure shares for five primary cereals (i.e., teff, wheat, barley, maize, and sorghum) and four food segments (i.e., cereals, pulses and root crops, fruit and vegetables, and animal products).[8] The detailed definitions of four food segments are presented in Table A1. A common problem in the

---

[5] The community questionnaires are administered by the field supervisors, rather than the enumerators. The price information are collected with the help of sellers in the markets.

[6] For the items that are missing in the community price data, the local prices are defined as the average of item-specific prices from the household expenditure data. For the items that are missing in both the community price data and the household expenditure data, the local prices are defined as the average prices from the three closest areas where such items are available.

[7] In the ESS, household consumption quantities comprise the quantities from the household purchases in last week, the stocks of self-production, and the gifts and other sources. The majority of household consumption comes from the first two sources.

[8] It worth noting that, due to the presence of zero observations, the average of expenditure conditional on non-zero observations is not equal to the product of the overall averages of price and quantity.

disaggregated data on food expenditure is that most expenditure variables are censored. In our context, there are large numbers of zeros in the data on cereal expenditures, since not every cereal is consumed in all EAs. The most consumed cereal is maize, consumed by 80 percent of EAs in 2013 and 85 percent of EAs in 2015. It is followed by teff, consumed by 80 percent of EAs in both 2013 and 2015. The problem of censored expenditures, however, is less of a concern for food segments since it is very likely that a few items, at least one, included in a food segment are purchased by some households in an EA.

During the survey period in 2013 and 2015, the prices of teff, wheat, and barley had a small variation, while the prices of maize and sorghum experienced a significant decrease. A food segment price is an average price weighted by sub-segment expenditure shares. Table 2.1 shows the prices of cereals and fruit and vegetables went down by 19 percent, while the prices of animal products increased slightly and the prices of pulses and root crops rose up by 53 percent. The changes in consumption quantities complied with the law of demand. In other words, the consumption quantities of cereals and fruit and vegetables increased, while the consumption quantities of pulses and root crops and animal products decreased. Lastly, the expenditure shares experienced minor changes between 2013 and 2015. Specifically, the expenditure shares of maize and sorghum decreased by 1.2 percent and 1.6 percent, whereas the expenditure shares of teff and wheat increased by 0.7 percent and 2.2 percent. On average, teff accounted for 40 percent of the food expenditure on cereals and the food segment of cereals accounted for a half of the total food expenditure. These suggest that any changes in cereal markets, especially in teff and wheat, can have important implications for Ethiopians' budgets and food security.

Table 2.2 displays the demographic statistics of sampled EAs from two waves of ESS in 2013 and 2015. There are 67 percent of EAs (i.e., 290 EAs) from the rural area, 10 percent from the small town area, and 23 percent from the urban area. The information on cereal production shows that maize is widely reported as a major crop grown in an EA, followed by teff and sorghum.

Particularly, there are 40 percent of EAs producing maize and 28 percent of EAs growing teff and sorghum as one of their major (or top three) crops in their communities. This is consistent with the information of the size of planting areas by cereal variety in Figure 2.1. There are 9 geographic regions included in the sample, in which chartered cities are defined as a separate group because they are significantly different from others in terms of the socioeconomic status. The geographic locations of EAs are presented in Figure A1.

2.4. Analytical Framework and Estimation Approach

To assess the impacts of alternative food policies on consumer welfare, we develop a two-level, structural demand system by following Chaudhuri, Goldberg, and Jia (2006). According to the theory of demand, a change in cereal prices will lead to changes in both the total cereal expenditure (i.e., the income effect) and the allocation of cereal consumption (i.e., the substitution effect). Therefore, the two-level demand system allows us to examine both the income and the substitution effect in the cereal segment due to the change in cereal prices resulting from alternative food-aid policies.

2.4.1. Demand

The food demand system is constructed in the approach of two-stage budgeting, according to the weak separability of household decisions on food expenditures.[9] In particular, households allocate their fixed total expenditures in two stages. At the upper stage, households divide the budget into several food segments (e.g., cereals, pulses and root crops, fruit and vegetables, and animal products), while at the lower stage, households decide the consumption of food items within each

[9] The weak separability implies that the food items can be partitioned into groups so that preferences within groups can be described independently of the quantities in other groups. For example, in our context, consumers can rank different bundles of cereals in a well-defined ordering regardless of their consumptions of items in other food segments.

segment. Food items included in the same segment, defined by the ESS (see Table A1), are assumed to be close substitutes.

First, consider the lower-level demand system, in which households allocate segment expenditures on food items within the segment. Specifically, let the relevant segment of cereals be indexed by G and the primary grains within the cereal segment be indexed by $i = 1, \dots, N$. Denote p_i and q_i, respectively, as the price and the quantity of grain i. The share of expenditure on grain i is ω_i, such that

$$\omega_i = \frac{p_i q_i}{\sum_j p_j q_j} = \frac{x_i}{X_G}, \tag{1}$$

where x_i is the expenditure on grain i and X_G is the expenditure on the cereal segment G. Following the specification of the Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980), the expenditure share equation is written as

$$\omega_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{X_G}{P_G}\right) + \epsilon_i, \tag{2}$$

where P_G is the price index given by

$$\ln P_G = \alpha_0 + \sum_i \alpha_i \ln p_i + \frac{1}{2}\sum_i \sum_j \tilde{\gamma}_{ij} \ln p_i \ln p_j, \tag{3}$$

and ϵ_i represents the additive error term for grain i. The properties of demand functions, derived from the utility maximization framework, suggest the following parameter restrictions, that is,

Adding-up: $$\sum_k \alpha_k = 1;\ \sum_k \beta_k = 0; \sum_k \tilde{\gamma}_{kj} = 0\ \forall j;$$

Homogeneity: $$\sum_k \tilde{\gamma}_{jk} = 0\ \forall j;$$

Symmetry: $$\gamma_{ij} = \gamma_{ji} = \frac{1}{2}\left(\tilde{\gamma}_{ij} + \tilde{\gamma}_{ji}\right).$$

The lower-level demand system is analyzed conditional on the segment expenditure (e.g., X_G). To address censoring in food expenditures on cereals, we adapt the form of the parsimonious censored system proposed by Shonkwiler and Yen (1999) for the lower-level demand system. Denote all parameters in the lower-level demand system, defined by (2), as θ and the explanatory part of the share equation for gain i as $f_i(\theta)$. Then the revised model with the observed expenditure share ω_i^o is given by

$$\omega_i^o = 1(Z_i\delta_i + \nu_i > 0)[f_i(\theta) + \epsilon_i], \tag{4}$$

where $1(\cdot)$ is a binary indicator function, Z_i is a vector of non-price and -income variables determining the likelihood of the demand for grain i with a conformable vector δ_i of parameters, and ν_i is the random error.

The upper-level demand system accounts for the changes in the segment expenditures due to the price changes in the cereal market. It is assumed that total food expenditures are spent on four segments, that is, cereals, pulses and root crops, fruit and vegetables, and animal products. Let the total expenditure X be given as a constant and the composite price index P be defined as a translog price index, analog to (3). The share of expenditure on segment G is ω_G, such that

$$\omega_G = \alpha_G + \sum_H \gamma_{GH} \ln P_H + \beta_G \ln\left(\frac{X}{P}\right) + \epsilon_G, \tag{5}$$

where all variables denoted by capital letters are defined as before but refer to segments rather than food items and the subscript H is a segment notation. The parameter restrictions are also imposed on the upper-level demand system.

The unconditional own- and cross-price elasticities are given by

$$\begin{aligned} \varepsilon_{ij} &= \frac{\partial \ln q_i}{\partial \ln p_j} = \frac{\partial \ln(\omega_i^o X_G / p_i)}{\partial \ln p_j} = \frac{\partial[\ln \omega_i^o + \ln X_G - \ln p_i]}{\partial \ln p_j} \\ &= \frac{\partial(\ln \omega_i^o - \ln p_i)}{\partial \ln p_j} + \frac{\partial \ln X_G}{\partial \ln p_j} \\ &= \varepsilon_{ij}\ |_{X_G = \bar{X}_G} + \frac{\partial \ln X_G}{\partial \ln P_G} \cdot \frac{\partial \ln P_G}{\partial \ln p_j} \end{aligned} \tag{6}$$

where the conditional price elasticities $\varepsilon_{ij}\ |_{X_G = \bar{X}_G}$ are such that

$$\varepsilon_{ij}\ |_{X_G=\bar{X}_G} = \frac{\partial \ln \omega_i^o}{\partial \ln p_j} - \delta_{ij}, \tag{7}$$

the Kronecker delta $\delta_{ij} = 1$ for $i = j$ and 0 for $i \neq j$ and the partial derivatives of $\ln \omega_i^o$ with respect to $\ln p_j$ are given by

$$\begin{aligned}
\frac{\partial \ln \omega_i^o}{\partial \ln p_j} &= \frac{\partial \omega_i^o}{\partial \ln p_j} \cdot \frac{1}{\omega_i^o} = \frac{\partial \omega_i}{\partial \ln p_j} \cdot \frac{1}{\omega_i} \\
&= \left[\gamma_{ij} + \beta_i \left(\frac{\partial \ln X_G}{\partial \ln p_j} - \frac{\partial \ln P_G}{\partial \ln p_j}\right)\right] \cdot \frac{1}{\omega_i} \\
&= \left[\gamma_{ij} + \beta_i \left(\frac{\partial \ln X_G}{\partial \ln P_G} - 1\right) \frac{\partial \ln P_G}{\partial \ln p_j}\right] \cdot \frac{1}{\omega_i} \\
&= \left[\gamma_{ij} + \beta_i \left(\frac{\partial \ln X_G}{\partial \ln P_G} - 1\right) \left(\alpha_j + \sum\nolimits_k \gamma_{kj} \ln p_k\right)\right] \cdot \frac{1}{\omega_i}.
\end{aligned} \tag{8}$$

By Chaudhuri, Goldberg, and Jia (2006), $\partial \ln X_G / \partial \ln P_G$ is well approximated by γ_{GG}/ω_G. Hence, we are able to calculate the unconditional and conditional price elasticities, defined in (6) and (7) respectively, with estimated parameters. Similarly, the unconditional income elasticities are given by

$$\begin{aligned}
\eta_i &= \frac{\partial \ln q_i}{\partial \ln X_G} = \frac{\partial \ln(\omega_i^o \omega_G X / p_i)}{\partial \ln X_G} = \frac{\partial [\ln \omega_i^o + \ln \omega_G + \ln X - \ln p_i]}{\partial \ln X_G} \\
&= \frac{\partial \ln \omega_i^o}{\partial \ln X_G} + \frac{\partial \ln \omega_G}{\partial \ln X} \frac{\partial \ln X}{\partial \ln X_G} + \frac{\partial \ln X}{\partial \ln X_G} \\
&= \frac{\partial \omega_i^o}{\partial \ln X_G} \cdot \frac{1}{\omega_i^o} + \left(\frac{\partial \omega_G}{\partial \ln X} \cdot \frac{1}{\omega_G} + 1\right) \cdot \frac{\partial \ln X}{\partial \ln X_G} \\
&= \frac{\beta_i}{\omega_i} + \left(\frac{\beta_G}{\omega_G} + 1\right) \cdot \frac{\partial X}{\partial X_G} \cdot \frac{X_G}{X} = \frac{\beta_i}{\omega_i} + \left(\frac{\beta_G}{\omega_G} + 1\right) \cdot \omega_G.
\end{aligned} \tag{9}$$

2.4.2. Estimation Approach

Ideally, we would like to have the aggregated data with the consumption information (i.e., expenditures and prices) of every food item for all EAs, so that we could follow the estimation approach of the two-level demand system proposed by Chaudhuri, Goldberg, and Jia (2006). However, we are not able to completely adopt their approach due to censoring in food expenditures on cereals and have to accommodate this issue in the estimation at the lower stage.

There are two data generating processes that can result in censored cereal expenditures. First, the data on cereal consumption are not well measured. By construction, the cereal consumption in an EA is the aggregated consumption overall households. Since the survey is nationally representative and households are well sampled, there is little chance that people in an EA do consume one cereal while all sampled households do not. So, this concern is less of a problem. Second, the data are well measured and people in an EA do not consume a certain cereal. In other words, there is no measurement error in the data. The zero cereal consumptions in an EA is attributed to the fact that there are no households consuming such cereals in this EA. The censored data in this study result from the second data generating process. In order to explain the zero cereal consumptions in an EA, we add in the lower-level demand a selection mechanism that determines the consumption of every cereal in the EA.

The selection mechanism is represented by as a binary choice model. Specifically, we use a probit model with a collection of variables capturing the fixed effects of regional lifestyles and socioeconomic circumstances and the supply shifters. The binary variables characterizing the fixed effects of regional lifestyles and socioeconomic circumstances are included in both the selection mechanism and the share equations of demand, whereas the supply shifters are only included in the selection model to address censoring in food expenditures on cereals. The supply shifters represent basic situations of the cereal production. In particular, the shifter variables indicate if primary cereals (i.e., teff, wheat, barley, maize, and sorghum) are grown as major crops (i.e., top three most grown crops) in an EA. These variables are predetermined and imply the market propensity for some cereals rather than the consumption quantity. As a result, households in an EA where teff is grown as a major crop are more likely to consume teff than households in another EA where teff is not grown as a major crop. With the information relevant to the cereal production, we cannot infer the cereal expenditure in an EA. Therefore, these binary variables delineating the production of major crops can be used to explain the selection mechanism in an EA.

To estimate the lower-level demand with censoring in food expenditures on cereals, we follow the two-step procedure developed by Shonkwiler and Yen (1999). Based on the assumption of the bivariate normality of (ν_i, ϵ_i), the observed expenditure share ω_i^o is thought of as the unconditional mean of the expenditure share, which is written as

$$\omega_i^o = \Phi(Z_i\delta_i)\left[f_i(\theta) + \lambda_i \frac{\phi(Z_i\delta_i)}{\Phi(Z_i\delta_i)}\right] + \xi_i = \Phi(Z_i\delta_i)f_i(\theta) + \lambda_i\phi(Z_i\delta_i) + \xi_i, \tag{10}$$

where $\xi_i = \omega_i^o - \mathrm{E}[\omega_i]$ is the disturbance term reflecting the error between the observed expenditure share and the unconditional, expected expenditure share, $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function (cdf) and the probability density function (pdf) of the univariate normal distribution, and $\phi(\cdot)/\Phi(\cdot)$ is the inverse mills ratio addressing the selection bias due to censoring. The scalar parameter λ_i represents the covariance of error terms (ν_i, ϵ_i) of each pair of the selection mechanism and the expenditure share of demand for cereal i. The estimation at the lower stage is conducted as follows: first, obtain ML estimates $\hat{\delta}_i$ of δ_i using the probit model for each cereal i; second, estimate the censored demand system for the cereal segment, defined by the $(n-1)$-equation system in (10) evaluated at $\{\hat{\delta}_i\}_{\forall i}$. In Deaton and Muellbauer's AIDS model, the homogeneity and symmetry restrictions are imposed in the estimation, while the adding-up restriction is implicitly built-in by dropping one of the share equations from this n-equation system. However, the adding-up restriction does not hold in the presence of censoring. Hence, in the second-step, we alternatively exclude one of n equations in each round of the estimation and the parameter estimates are the means of $(n-1)$ associated estimates from all n alternative estimations. Subsequently, the price and income elasticities in (6) and (9) are respectively obtained. All of the standard errors of parameter and elasticity estimates are calculated by using a bootstrap technique (also see in Bilgic and Yen 2013).[10]

[10] We use a nonparametric bootstrap technique (i.e., Monte Carlo simulation), in which our data are treated as the population. There are 500 bootstrap samples of size N (= 865) drawn from our data with replacement.

2.5. Results

2.5.1. Parameter Estimates and Elasticities

Tables A2 and A3 in the Appendix display the estimates of parameters in the selection equation and the lower-level demand system. In Table A2, the indicators of major crops characterizing basic situations of the cereal production are positively related to the decision on the cereal consumption. In Table A3, the coefficients of inverse mills ratios are significant, suggesting the correlations between the consumption decisions on cereals and the associated level demand. In other words, it justifies the inclusion of the selection mechanism at the lower stage. Table A4 shows the results of the demand estimation at the upper stage. The coefficient of year dummy is either small in the segment share equation of animal products or insignificant in other segment share equations. This finding implies that the consumption pattern was similar between the two waves. For ease of interpretation, we show the impacts of explanatory variables in the selection equation on cereal consumption decisions in Table 2.3 and present the unconditional price and income elasticities at the upper and the lower stage in Table 2.4 and 2.5, respectively.

In Table 2.3, the marginal effects of regional dummies suggest the cereal consumption patterns vary across regions. Also, the consumption patterns are significantly different between rural and urban (or small town) areas. The likelihoods of consuming teff in a small town EA and an urban EA are 35.2 percent and 43.8 percent higher than that in a rural area, respectively. Moreover, we find the indicators of major crops play important roles in the selection mechanism. Compared to an EA where teff is not a major crop, an EA where teff is a major crop is 18 percent higher in the propensity for consuming teff. This difference is even greater in terms of sorghum.

One set of the second-step estimates (i.e., θ and λ) can be obtained from the estimation using a bootstrap sample, so can be the elasticity estimates (i.e., ε and η). All the standard errors are the standard deviations over the associated estimates obtained from 500 bootstrap samples.

Specifically, sorghum is 55.2 percent more likely to be consumed in an EA where it is a major crop than another EA where it is not.

Table 2.4 and 2.5 present price and income elasticities at the upper and lower stage, respectively. The diagonals of Marshallian and Hicksian elasticity matrices include the estimates of the own-price elasticities, all of which are negative and significant. As expected, the estimates of Hicksian own-price elasticities are all smaller than the estimates of Marshallian ones, since foods are normal goods in Ethiopia. Table 2.4 shows the consumption pattern across food segments. The estimates of Marshallian own-price elasticities are less than one in magnitude, implying cereals, pulses and root crops, fruit and vegetables, as well as animal products are staple foods for Ethiopian households. The estimates of Marshallian cross-price elasticities show that these foods are generally complementary goods. The magnitude of these estimates implies there are little complementary effects among these foods from different segments, consistent with the weak separability assumption. The estimates of Hicksian price elasticities reflect the substitution pattern across food segments with a given utility level. For example, to remain at the same utility level, a 1 percent decrease in the price of cereals will lead to an increase of 0.73 percent in the consumption of animal products. The estimates of income elasticities suggest, among these food segments, animal products are the upper-class foods for Ethiopian households. An increase of 10 percent in household expenditure on foods will result in an increase of 14 percent in the consumption of animal products.

Table 2.5 displays the detail substitution and income effects within the cereal segment, which provides insights for the subsequent analysis of consumer welfare. The income effects of cereals are all positive and significant, suggesting all cereals are normal goods. In particular, every 10 percent increase in food expenditure on cereals will raise the consumption of teff—the most favorable cereal—by 6.33 percent, wheat by 6.29 percent, and maize—the least one—by 1.07 percent. The income elasticity of barley and sorghum are in between. As the utility level is fixed in the subsequent analysis of consumer welfare, we focus our discussion thereafter on Hicksian elasticities in this section and the term "Hicksian" is suppressed. The own-elasticities are negative ranging from –0.39 to –0.87, which strongly suggest all the cereals are inelastic with respect to price and considered as staple foods for Ethiopian households.

Among all of these primary cereals, teff is the most inelastic, implying its essential role in Ethiopian daily diets. The cross-price elasticities of teff with respect to other cereal prices are positive and highly significant, which means that other cereals are close substitutes for teff. In other words, the consumption of teff will decline as the prices of wheat, barley, maize, and sorghum fall down. For example, the consumption of teff will drop by 1.57 percent if the price of wheat decreases by 10 percent and 2.47 percent if the price of maize decrease by 10 percent. These significant cross-price elasticities of teff to competing cereals inform the potential policies on food aid. In particular, consumer welfare can be improved if we think of facilitating a consumption switch from teff to other less expensive cereals in response to the increasing teff prices.

The cross-price elasticities of wheat indicate that barley is a complement, while teff, maize, and sorghum are substitutes. Similarly, maize and sorghum are complements, but both of them are substitutes for teff, wheat, and barley. The magnitude of the cross-price elasticities of wheat suggests wheat is more reactive to changes in the prices of other cereals than teff. For example, the consumption of wheat will increase by 3.69 percent as the price of maize increases by 10 percent and 4.22 percent as the price of sorghum increases by 10 percent. In addition, comparing the significance and magnitude of cross-price elasticities across columns in Table 2.5, we find the cereal market is more responsive to the changes in the prices of wheat and sorghum than barley and maize. This suggests that wheat and sorghum can be better instruments used to influence the equilibrium consumption of cereals in Ethiopia.

2.5.2. Welfare Analysis

In this section, we introduce how to measure the difference in consumer welfare as the prices change and conduct simulations under alternative food aid policies in response to the increasing teff prices. The simulation results provide insights into how the Ethiopian government and non-profit organizations ought to react to the increasing teff prices. To compare with the effectiveness of alternative policies on food aid, we compute the distance of every alternative policy to the ideal one using a measurement of dissimilarity.

We measure the change in consumer welfare by the compensating variation (CV)—the additional expenditure that consumers have to pay for living with a constant utility as teff prices hike up to a new level. Specifically, let $\boldsymbol{p}^0$ and $\boldsymbol{p}^1$ be the prices before and after the increase in teff prices, u^0 be the level of initial utility attained by consumers, and $E(\boldsymbol{p}, u^0)$ be the expenditure function. Then, the CV is given by

$$\mathrm{CV} = E(\boldsymbol{p}^1, u^0) - E(\boldsymbol{p}^0, u^0). \tag{11}$$

Note that there are two ways to calculate the value of CV. One is to directly compute the value of CV using $E(\boldsymbol{p}, u^0)$. The other is to attain the second order approximation of CV using the Taylor series expansion of $E(\boldsymbol{p}^1, u^0)$ around $\boldsymbol{p}^1 = \boldsymbol{p}^0$; that is,

$$\mathrm{CV} = (\boldsymbol{p}^1 - \boldsymbol{p}^0)^T \cdot \frac{\partial E(\boldsymbol{p}^0, u^0)}{\partial \boldsymbol{p}^0} + \frac{1}{2} \cdot (\boldsymbol{p}^1 - \boldsymbol{p}^0)^T \cdot \frac{\partial^2 E(\boldsymbol{p}^0, u^0)}{(\partial \boldsymbol{p}^0)^2} \cdot (\boldsymbol{p}^1 - \boldsymbol{p}^0). \tag{12}$$

Suppose all other cereal prices remain constant as the price of teff increases. The second order approximation of CV in (12) is then reduced to

$$\mathrm{CV} = \left[\frac{p_{\mathrm{teff}}^1 - p_{\mathrm{teff}}^0}{p_{\mathrm{teff}}^0} + \frac{1}{2}\varepsilon_{\mathrm{teff}}^h \left(\frac{p_{\mathrm{teff}}^1 - p_{\mathrm{teff}}^0}{p_{\mathrm{teff}}^0}\right)^2\right] \omega_{\mathrm{teff}}^0 X_G^0, \tag{13}$$

where $\varepsilon_{\mathrm{teff}}^h$ is the Hicksian own-price elasticity of teff, ω_{teff}^0 the expenditure share of teff within the cereal segment before the increase in teff prices, and X_G^0 is the original food expenditure on cereals.

We examine the changes in the weekly cereal consumption for an average household as teff prices increase under alternative policies. Figure 2.2 shows that there was about a 10 percent increase in teff prices from 2015 to 2016. Therefore, in Table 2.6, we present simulation results from the scenario in which teff prices increase by 10 percent. For more information, Table 5A in the Appendix displays the simulation results as teff prices increase by 5 percent and 15 percent. There are five alternative policy measures reacting to the increasing teff prices. The ideal food aid policy is to offer households cash voucher, as the same amount of CV, allowing them to achieve the same utility with the new, increased teff prices. In particular, using the expenditure function derived from the estimated demand system, we calculate the CV in (11) given a 10 percent increase in teff prices. The associated CV is 7.82 Birrs a week, that is, 3.85 percent of the weekly expenditure on foods for an average household.[11] In Table 2.5, the simulation results from the estimated demand system shows that a cash voucher of 7.82 Birrs enables an average household to maintain a constant utility by increasing their consumptions of wheat from 2.97 to 3.14 kilograms, barley from 1.02 to 1.26, maize from 3.76 to 6.09, and sorghum from 2.51 to 3.32, while decreasing their consumptions of teff from 4.21 to 4.06 kilograms. Now, suppose instead of cash vouchers the government launched a food aid policy to provide households with wheat. Then, an average household will receive 1.09 (= 4.07 – 2.98) kilogram of wheat, worth 7.82 Birrs, as food aids. Similarly, for a food aid providing barley, maize, and sorghum, an average household will receive 1.10 (= 2.31 – 1.21) kilogram of barley, 1.88 (= 7.81 – 5.93) kilograms of maize, and 1.44 (= 4.63 – 3.19) kilograms of sorghum.

As a robustness check, we also examine the changes in the cereal consumption using Hicksian and Marshallian price elasticities. The estimate of CV given by (13) is 4.55 Birrs. The cereal consumption associated with the food aid of offering cash voucher is calculated using

[11] According to data, the weekly expenditure on foods for an average household is 203 Birrs.

Hicksian price elasticities and the cereal consumption from the data. Table 2.6 reveals that, given a cash voucher of 4.55 Birrs, an average household will decrease teff and maize consumption to 4.04 and 3.71 kilograms, whereas increase wheat, barley, and sorghum consumption to 3.06, 1.05, and 2.53 kilograms. Then we compute changes in the cereal consumption for an average household with a fixed total expenditure on food using Marshallian price elasticities and obtain the policy outcomes of cereal consumption by adding the associated cereal at the same amount of cash voucher worth 4.55 Birrs. The additional consumption of wheat, barley, maize, and sorghum associated with the corresponding food aids are 0.64, 0.64, 1.09, and 0.84 kilogram(s), respectively. Compared to the cereal consumption simulated by the estimated demand system, the results based on the elasticities are smaller, suggesting the non-linearity of the food expenditure function.[12]

Next, we evaluate the distances between the ideal food aid policy of offering cash voucher and the alternative food aid policies of providing real cereals. To measure the relative dissimilarity in terms of quantity between two consumption bundles, we adopt the weighted log quadratic formula given by (Diewert 2009), that is,

$$\Delta(\boldsymbol{p}^0, \boldsymbol{p}^1, \boldsymbol{q}^0, \boldsymbol{q}^1) = \sum_{i=1}^{n} \frac{\omega_i^1 + \omega_i^0}{2} \left[\ln\left(\frac{q_i^1}{q_i^0 Q_{01}} \right) \right]^2, \tag{14}$$

where $\boldsymbol{q}^0$ and $\boldsymbol{q}^1$ are the vectors of consumption quantities before and after the increase in teff prices. The scalar Q_{01} denotes the Törnqvist index, given by

$$Q_{01} = \prod_{i=1}^{n} \left(\frac{q_i^1}{q_i^0} \right)^{(\omega_i^1 + \omega_i^0)/2}. \tag{15}$$

The last columns of Table 2.6 and Table A5 present the measure of dissimilarity toward the ideal food aid policy of offering cash voucher. The results from both simulations using the estimated demand system and elasticities in Table 2.6 indicate that the policy offering wheat as food aid is

---

[12] The simulation results of CV using the estimated demand system and elasticities are closer to each other as the increase in teff prices is smaller (see Table 5A).

the option closest to the ideal alternative—offering cash voucher. This finding is robust across scenarios where teff prices increase by 5 and 15 percent (see Table A5).

2.6. Conclusion

Teff is an ancient grain primarily grown in Ethiopia and has been a staple food for Ethiopians for millennia. As a popular ingredient in contemporary healthy diets in developed countries, teff has experienced an increase in both prices and demand in the international market. This "teff fever" inevitably affects consumer welfare in Ethiopia especially after the Ethiopian government withdrew the export restriction on teff in mid-2015.

This chapter examines the changes in consumer welfare due to the increasing teff prices and the roles of multiple food aid policies in reducing the adverse impacts of the increasing teff prices by estimating a two-stage food demand system with a focus on cereals. In the estimation, we accommodate censoring in food expenditures on cereals by adapting a censored demand system at the lower stage. A selection mechanism is introduced to determine the propensity of cereal consumption in an EA. We find that the indicators of major crops that characterize the situation of cereal production are positively associated with the propensity of cereal consumption.

Using the full estimated demand system, we calculate the price and income elasticities of cereals. Although the total expenditures on food staples might change as the Ethiopian economy grows, the consumption patterns featured by elasticities would remain fairly stable. The cereal price and income elasticities inform policymakers about how consumers respond to the changes in prices and their incomes. We find that own-price elasticities of all primary cereals are negative and significant, and teff is the most inelastic one among all primary cereals. This implies that cereals, especially teff, are staple foods for Ethiopian households. Therefore, changes in teff prices might have profound impacts on consumer welfare. Moreover, the cereal market is more responsive to the changes in the prices of wheat and sorghum than barley and maize. This implies that wheat and sorghum have better potential to be used as instruments for food aid policies.

Finally, we evaluate the changes in consumer welfare as teff prices increase and simulate the weekly consumption of all cereals under alternative food aid policies for an average household. Specifically, given a 10 percent increase in teff prices, the monetary compensation for living in a constant level of utility—the value of CV—is 7.82 Birrs a household a week evaluated by the estimated demand system and 4.55 Birrs by the second order approximation using elasticities, accounting for 3.85 percent and 2.24 percent of the weekly expenditure on foods for an average household, respectively. This discrepancy implies the non-linearity of the food expenditure function. The smaller a change in teff prices is, the smaller this discrepancy will be. In addition to the ideal food aid policy that offers households cash voucher, we discuss alternative policies that provide households real cereals at the same value of the cash voucher. This discussion is informative and practical since a poor developing country usually receives real cereals rather than cash vouchers as food aid from non-profit organizations and other countries. The simulation results show that the policy providing wheat as food aid is the most effective alternative close to the cash voucher solution. In other words, compared to maize, sorghum, and barley, the outcomes of the policy offering wheat as food aid are closer to the outcomes of the policy offering cash voucher. This finding lends support to the ongoing Ethiopian food policy that distributes subsidized wheat on a large scale.

Table 2.1. Sample Statistics of Expenditures, Prices, Quantities, and Expenditure Shares

	2013 (N = 433)			2015 (N = 432)		
	% Consuming	Mean	SD	% Consuming	Mean	SD
Expenditures (million Birrs)						
Cereals	100.00	4.37	6.79	100.00	4.78	6.58
Teff	79.68	1.59	3.14	79.63	2.12	3.87
Wheat	78.98	0.75	1.81	78.01	0.98	2.39
Barley	62.36	0.46	2.37	60.19	0.33	1.20
Maize	79.91	0.90	3.04	84.95	0.72	1.31
Sorghum	62.12	0.67	2.27	62.27	0.63	1.82
Pulses and Root Crops	96.77	1.59	3.72	97.69	2.09	6.20
Fruit and Vegetables	98.15	0.41	0.54	97.92	0.47	0.57
Animal Products	96.77	1.83	3.41	97.22	2.01	4.86
Prices (Birrs per kilogram)						
Cereals		9.01	8.46		8.09	4.34
Teff		11.00	5.26		11.45	3.70
Wheat		7.44	3.81		7.45	2.63
Barley		7.46	6.26		7.53	4.38
Maize		5.52	6.78		4.46	2.39
Sorghum		6.42	6.39		5.47	3.31
Pulses and Root Crops		9.96	5.05		15.21	10.67
Fruit and Vegetables		9.24	4.49		7.50	3.27
Animal Products		43.64	26.21		44.26	26.23
Quantities (thousand kilograms)						
Cereals		629.05	903.07		771.15	1150.87
Teff		153.38	311.95		200.66	380.75

Wheat	116.75	278.99	163.74	521.90
Barley	65.40	244.97	65.51	257.69
Maize	173.06	371.42	210.39	403.09
Sorghum	120.46	329.98	130.85	357.18
Pulses and Root Crops	256.86	663.08	214.91	438.58
Fruit and Vegetables	51.33	72.14	70.74	83.82
Animal Products	77.10	121.65	77.00	137.01
Expenditure shares (percent)				
Cereals	52.68	20.16	51.60	20.36
Teff	38.28	32.55	39.02	31.55
Wheat	16.32	18.83	18.59	21.51
Barley	6.45	13.88	6.22	12.72
Maize	21.48	25.68	20.26	23.20
Sorghum	17.47	25.35	15.92	22.91
Pulses and Root Crops	16.88	14.16	19.68	14.23
Fruit and Vegetables	6.73	5.16	6.38	5.18
Animal Products	23.71	18.05	22.34	17.59

Note: All statistics are calculated at the EA level. The column labeled “% Consuming” presents the percentage of non-zero observations in the sample. The column labeled “Mean” (“SD”) presents the conditional average (standard deviation) of expenditure over non-zero observations, and the unconditional average (standard deviation) of price, quantity, and expenditure share. The food segment of cereals is made of teff, wheat, barley, maize, and sorghum. The expenditure shares of teff, wheat, barley, maize, and sorghum sum to 1, and the expenditures of cereals, pulses and root crops, fruit and vegetables, and animal products sum to 1. All prices and expenditures are deflated by the CSA CPI for food and non-alcoholic beverages with the base period in November 2011.

Table 2.2. Definitions and Sample Statistics of Demographic Variables (N = 865)

Variable	Definition	Mean	SD
Year 2015	Survey in 2015	0.50	0.50
Rural	Rural area (reference)	0.67	0.47
Town	Small town area	0.10	0.30
Urban	Urban area	0.23	0.42
Major crop - teff	Teff is one of the major (top 3) crops grown in this enumeration area	0.28	0.45
Major crop - wheat	Wheat is one of the major (top 3) crops grown in this enumeration area	0.20	0.40
Major crop - barley	Barley is one of the major (top 3) crops grown in this enumeration area	0.14	0.34
Major crop - maize	Maize is one of the major (top 3) crops grown in this enumeration area	0.40	0.49
Major crop - sorghum	Sorghum is one of the major (top 3) crops grown in this enumeration area	0.28	0.45
Region - Tigray	Enumeration area is in Tigray (reference)	0.11	0.32
Region - Afar	Enumeration area is in Afar	0.03	0.17
Region - Amhara	Enumeration area is in Amhara	0.20	0.40
Region - Oromiya	Enumeration area is in Oromiya	0.20	0.40
Region - Somali	Enumeration area is in Somali	0.06	0.24
Region – Bens. Gumuz	Enumeration area is in Benshangul Gumuz	0.03	0.16
Region - SNNP	Enumeration area is in SNNP	0.23	0.42
Region - Gambella	Enumeration area is in Gambella	0.03	0.16
Region - Chartered Cities	Enumeration area is in Harari, Addis Ababa, and Dire Dawa	0.12	0.32

Table 2.3. Marginal Effects of Explanatory Variables in the Selection Equation

	Teff	Wheat	Barley	Maize	Sorghum
Year 2015	0.012	0.008	0.005	0.065***	0.028
	(0.021)	(0.025)	(0.029)	(0.023)	(0.029)
Region - Afar	–0.114*	0.209**	–0.214**	0.132	–0.453***
	(0.062)	(0.082)	(0.094)	(0.084)	(0.100)
Region - Amhara	–0.050	0.020	0.023	–0.033	–0.123**
	(0.047)	(0.046)	(0.054)	(0.035)	(0.054)
Region - Oromiya	–0.071	0.095**	0.171***	0.085**	–0.079
	(0.046)	(0.047)	(0.055)	(0.039)	(0.054)
Region - Somali	–0.528***	0.141**	–0.292***	0.010	–0.227***
	(0.065)	(0.061)	(0.076)	(0.059)	(0.075)
Region - Bens. Gumuz	0.026	–0.095	–0.188*	0.031	0.097
	(0.071)	(0.072)	(0.102)	(0.104)	(0.147)
Region - SNNP	–0.068	0.020	0.127**	0.236***	–0.076
	(0.044)	(0.044)	(0.053)	(0.047)	(0.053)
Region - Gambella	–0.120*	–0.220***	–0.043	0.093	–0.409***
	(0.062)	(0.072)	(0.093)	(0.114)	(0.101)
Region - Chartered Cities	–0.218***	0.210***	–0.104*	–0.049	–0.032
	(0.050)	(0.060)	(0.061)	(0.040)	(0.065)
Town	0.352***	0.146***	0.129***	0.054	0.074
	(0.077)	(0.042)	(0.050)	(0.041)	(0.048)
Urban	0.438***	0.250***	0.299***	0.009	–0.125***
	(0.055)	(0.036)	(0.035)	(0.028)	(0.036)
Major crop – teff	0.180***				
	(0.027)				
Major crop – wheat		0.381***			
		(0.049)			
Major crop – barley			0.446***		
			(0.055)		
Major crop – maize				0.271***	
				(0.034)	
Major crop – sorghum					0.552***
					(0.047)

Note: Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table 2.4. Demand Pattern across Food Segments: Price and Income Elasticities

	Cereals	Pulses and Root Crops	Fruit and Vegetables	Animal Products
Marshallian price elasticities				
Cereals	–0.9234***	–0.0128**	–0.0133***	–0.0290***
	(0.0077)	(0.0057)	(0.0027)	(0.0047)
Pulses and Root Crops	–0.0056**	–0.9899***	–0.0008	–0.0022
	(0.0023)	(0.0025)	(0.0007)	(0.0014)
Fruit and Vegetables	–0.0005*	0.0001	–0.9980***	0.0003***
	(0.0003)	(0.0002)	(0.0004)	(0.0001)
Animal Products	–0.0211***	–0.0054***	–0.0014***	–0.9881***
	(0.0024)	(0.0014)	(0.0005)	(0.0016)
Hicksian price elasticities				
Cereals	–0.426***	0.178***	0.040***	0.156***
	(0.019)	(0.011)	(0.003)	(0.008)
Pulses and Root Crops	0.513***	–0.791***	0.054***	0.190***
	(0.028)	(0.015)	(0.003)	(0.012)
Fruit and Vegetables	0.230***	0.088***	–0.974***	0.086***
	(0.030)	(0.012)	(0.003)	(0.011)
Animal Products	0.732***	0.283***	0.079***	–0.709***
	(0.041)	(0.018)	(0.005)	(0.018)
Income elasticities				
	0.926***	0.964***	0.428***	1.401***
	(0.028)	(0.053)	(0.054)	(0.075)

Note: Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table 2.5. Demand Pattern within the Cereal Segment: Price and Income Elasticities

	Teff	Wheat	Barley	Maize	Sorghum
Marshallian price elasticities					
Teff	–0.632***	0.031	0.144***	0.090*	0.085
	(0.065)	(0.051)	(0.055)	(0.051)	(0.052)
Wheat	0.038	–0.651***	–0.181***	0.213**	0.311***
	(0.082)	(0.109)	(0.064)	(0.088)	(0.080)
Barley	0.218	–0.529***	–0.642**	0.116	0.037
	(0.191)	(0.140)	(0.278)	(0.170)	(0.156)
Maize	–0.130*	0.034	0.008	–0.895***	–0.316***
	(0.069)	(0.070)	(0.069)	(0.126)	(0.085)
Sorghum	–0.029	0.195***	0.034	–0.247***	–0.716***
	(0.072)	(0.068)	(0.063)	(0.083)	(0.127)
Hicksian price elasticities					
Teff	–0.390***	0.157***	0.195***	0.247***	0.197***
	(0.061)	(0.054)	(0.055)	(0.059)	(0.055)
Wheat	0.278***	–0.525***	–0.131**	0.369***	0.422***
	(0.085)	(0.111)	(0.065)	(0.099)	(0.087)
Barley	0.366*	–0.451***	–0.611**	0.212	0.105
	(0.190)	(0.156)	(0.288)	(0.191)	(0.179)
Maize	–0.091	0.055	0.017	–0.867***	–0.296***
	(0.072)	(0.072)	(0.071)	(0.130)	(0.090)
Sorghum	0.114	0.270***	0.064	–0.153	–0.649***
	(0.078)	(0.070)	(0.066)	(0.094)	(0.136)
Income elasticities					
	0.633***	0.629***	0.387**	0.107*	0.375***
	(0.042)	(0.056)	(0.169)	(0.057)	(0.072)

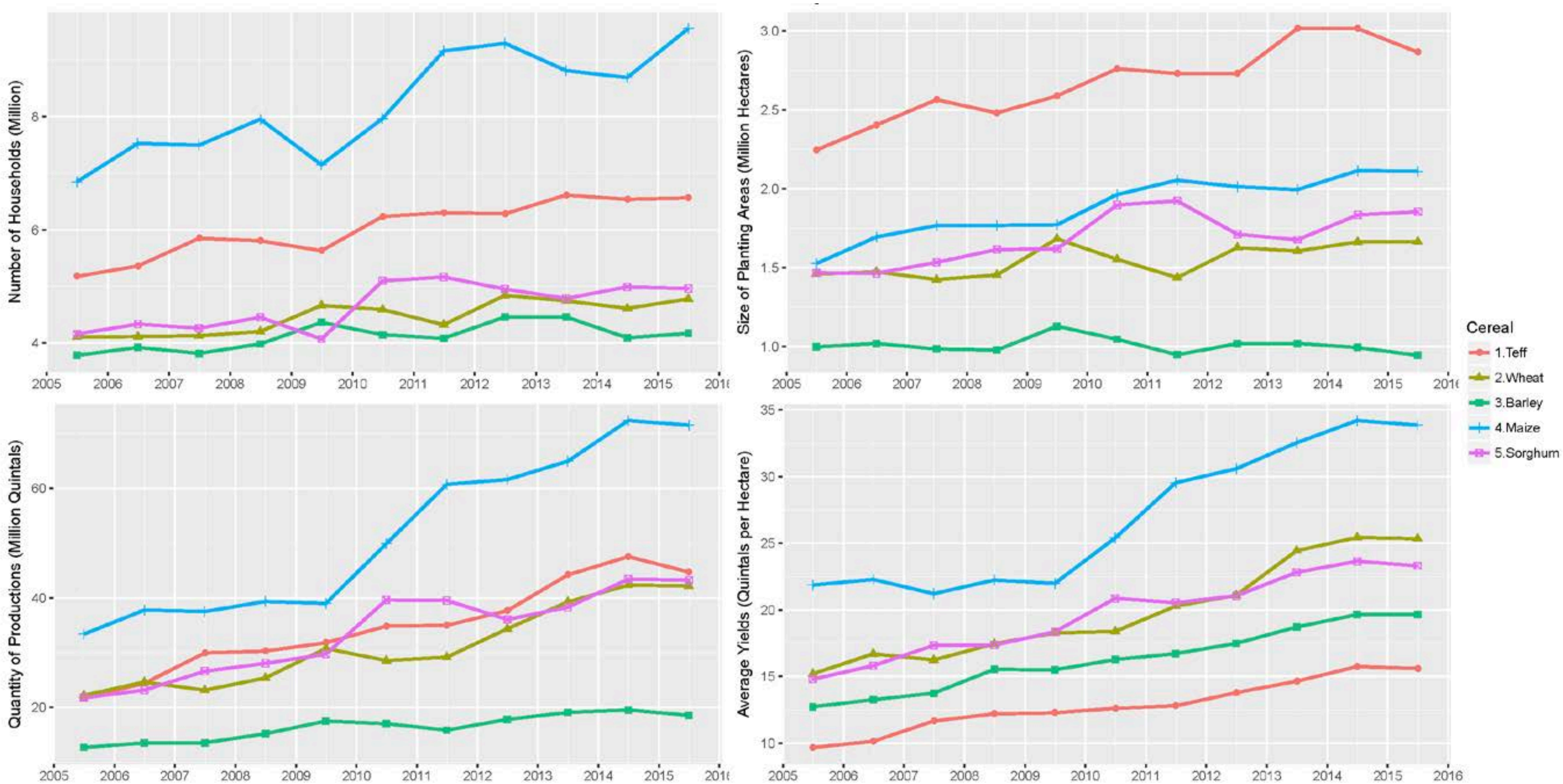
Note: Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table 2.6. Policy Outcomes of Weekly Cereal Consumption as Teff Prices Increase by 10 Percent (Kilogram)

Policy	Teff	Wheat	Barley	Maize	Sorghum	Dissimilarity from Cash Voucher
Cereal consumptions for an average household a week from the ESS 2015/2016						
	4.21	2.97	1.02	3.76	2.51	
Cereal consumptions for an average household a week simulated using estimated demand system as teff prices increase by 10 percent with CV as 7.82 Birrs						
Food aid – cash voucher	4.06	3.14	1.26	6.09	3.32	0.00
Food aid – wheat	3.85	4.07	1.21	5.93	3.19	1.50
Food aid – barley	3.85	2.98	2.31	5.93	3.19	3.86
Food aid – maize	3.85	2.98	1.21	7.81	3.19	1.57
Food aid – sorghum	3.85	2.98	1.21	5.93	4.63	2.04
Cereal consumptions for an average household a week simulated using estimated elasticities as teff prices increase by 10 percent with CV as 4.55 Birrs						
Food aid – cash voucher	4.04	3.06	1.05	3.71	2.53	0.00
Food aid – wheat	3.94	3.62	1.04	3.70	2.50	0.61
Food aid – barley	3.94	2.98	1.68	3.70	2.50	1.93
Food aid – maize	3.94	2.98	1.04	4.79	2.50	1.06
Food aid – sorghum	3.94	2.98	1.04	3.70	3.34	1.10

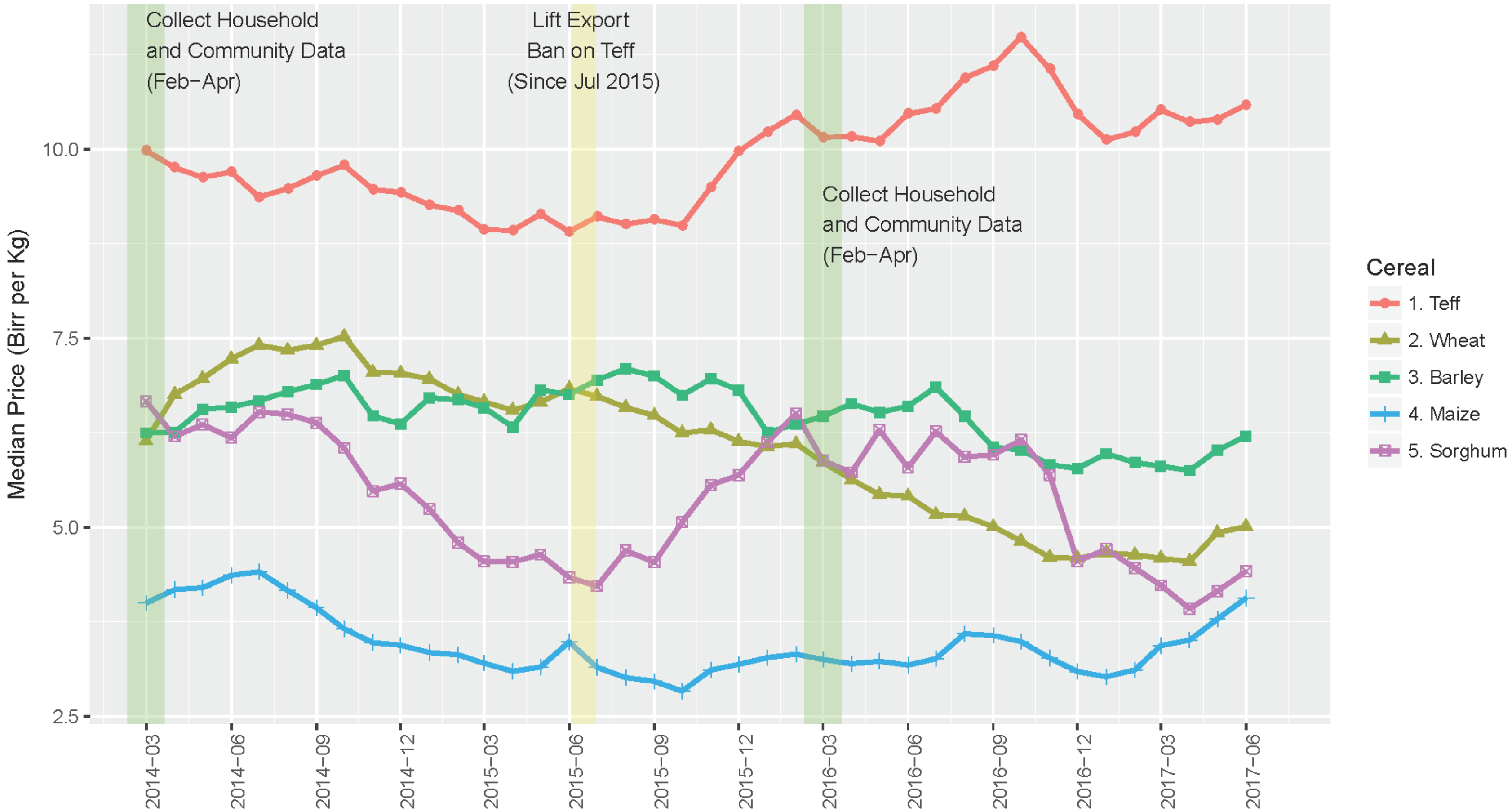
Note: The first part shows the weekly cereal consumptions for an average household at the expenditure of 203 Birrs from data. To examine changes in household cereal consumptions a week, the second part presents the simulation results from the estimated demand system while the third part presents the simulation results using Marshallian and Hicksian elasticities. According to Equation 14, the last column shows the value of dissimilarity measure ($\times$ 100) between food aid policy offering cash–voucher and that providing wheat only, barley only, maize only, and sorghum only.

Figure 2.1. Households, Area, Production, and Yield for Primary Cereals in Meher Season



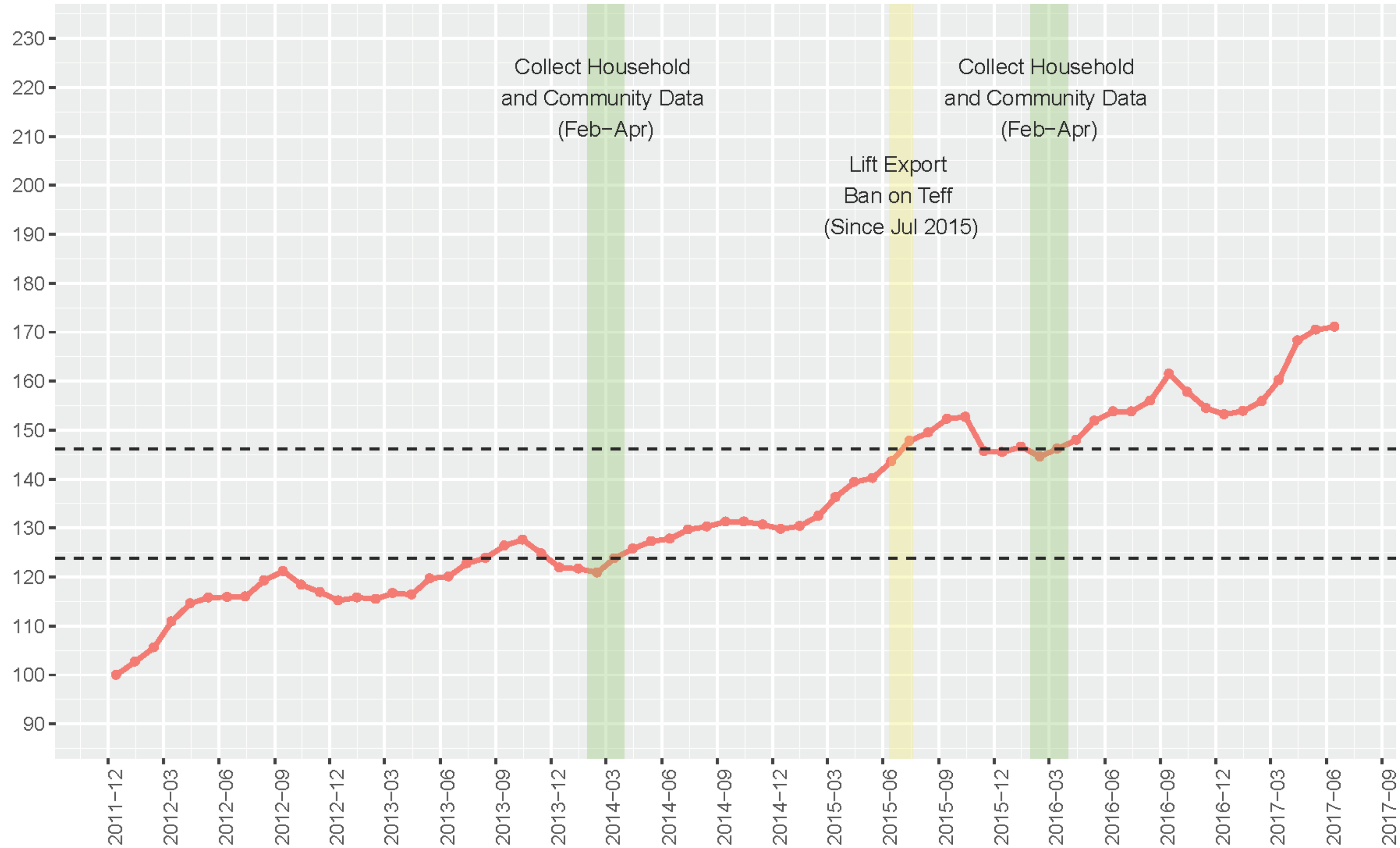
Note: Data come from the Central Statistical Agency annual reports on the Agricultural Sample Survey in Ethiopia from 2006 to 2016. The Meher season is the main crop season and produces 90 to 95 percent of total grain output (USDA 2008).

Figure 2.2. Wholesale Prices for Primary Cereals in Ethiopia



Note: Data come from the market statistics reported by the Ethiopian Grain Trade Enterprise.

Figure 2.3. Consumer Price Index in Ethiopia



Note: Data are from the Central Statistical Agency in Ethiopia with the base period in December 2011. The dash lines indicate the levels of CPI for food and non–alcoholic beverages when household and community data were collected in the Ethiopia Socioeconomic Survey 2013–2014 and 2015–2016 respectively.

Chapter 3. The Welfare Impacts of New Demand-Enhancing Agricultural Products

3.1. Introduction

Over the past decade, food and agricultural markets have become more consumer-oriented (Unnevehr et al. 2010). Large numbers of new products are developed to serve consumers' heterogeneous tastes and increasing expectations of food quality. According to Mintel's Global New Product Database, the agriculture sector in the United States (U.S.) has introduced to markets more than 3,500 new varieties of fruit and vegetables since 2011 (USDA 2017a). However, little is known about the magnitude of consumer benefits from the introduction of demand-enhancing agricultural products.

Consumers are affected by the introduction of a new agricultural product in two ways. First, some consumers are better off with a growing number of varieties because the new variety with different attributes might better serve their preferences.[13] This is interpreted as an impact mainly capturing the "consumer preference for diversity." Second, consumers will directly receive an economic benefit if the new variety increases market competition and leads to lower prices of other varieties. These lower prices would then attract more consumers and hence increase the aggregate demand for all varieties.

In this chapter, we estimate the change in consumer welfare from the introduction of a new variety, the Honeycrisp, in the U.S. apple market. The apple market has several desirable features for the purposes of this study. First, apples are the second most valuable fruit in the United States (USDA 2016b) and the fruit is marketed by variety names or trademarked brand names. Second, the market is also very dynamic with a large number of newly patented varieties under development (Rickard et al. 2013). Third, the findings have important implications for public

[13] The terms "attributes" and "characteristics" will be used interchangeably through this chapter.

investments as the growth of the apple industry is rooted in the success of the breeding programs at land-grant universities such as Cornell University, Washington State University, and the University of Minnesota.

Public organizations and the U.S. government have a long history of funding agricultural research and development (R&D) programs through the university systems (Foltz, Barham, and Kim 2000). However, the growth rate of public investment in agricultural R&D began to decrease in the early 1950s. By 1974, more than half of the total investments were provided by the private sector, and this ratio increased to 58% in 2009 (Pardey et al. 2015). The development of patent protection laws mitigates the adverse effects of the decline of public support and encourages more private investments in agricultural R&D (Pray and Fuglie 2015). The Bayh-Dole Act of 1980 and the subsequent legislation gave universities the permission to attain the ownership of inventions made with federal funding and, thereby, enabled them to finance agriculture research by transferring a part of their patent rights to the private sector.[14]

Particularly, we evaluate the welfare changes in the apple market from the introduction of Honeycrisp apples using structural models of consumer demand and retailer competition. On the demand side, we estimate a random utility model of demand that explicitly accounts for consumers' heterogeneous tastes and preferences. On the supply side, we model the retailer competition in a Bertrand-Nash fashion and derive the pricing rules for apples. Using the estimated demand parameters together with the pricing rules, we consequently simulate equilibrium outcomes in a counterfactual scenario wherein Honeycrisp apples are removed from the market. Then we quantify the changes in consumer welfare, market size, and sales revenue. We obtain data from multiple sources. The primary data are the point-of-sale scanner data which include apple prices and sales revenues at the Universal Product Code (UPC) level from 61 cities across the United States in the

[14] The provisions of the Bayh-Dole Act were further supported by the Federal Technology Transfer Act of 1986 and the National Technology Transfer and Advancement Act of 1996.

period from March 2009 to February 2015. The rest of our data comprise the population statistics of demographics, such as age and household income, and the cost data for retailers, such as apple prices in the wholesale market and wage rates in the retailing industry.

The results show that the introduction of Honeycrisp apples drives the prices of competing apple varieties downward, especially for the best-selling varieties such as Gala and Red Delicious. The extent of decline in prices is positively correlated with the market share of Honeycrisp apples. On average, the prices of Gala and Red Delicious apples decrease by 0.72 percent and 0.61 percent respectively, when the market share of Honeycrisp apples is greater than or equal to 1 percent. If the share rises up to 5 percent, the prices of Gala and Red Delicious apples decrease by 2.23 percent and 1.67 percent respectively. Compared to the results in the counterfactual scenario wherein Honeycrisp apples are removed from the markets, the estimates show that Honeycrisp has increased the total sales quantity by 8.03 percent and the total sales revenue by 21.25 percent over the study period. In addition, the results show that the consumer welfare increases from 3.03 million dollars in 2009 to 15.20 million dollars in 2014. The total changes in consumer welfare can be decomposed into the changes due to the increased number of total apple varieties and the changes due to the decline in prices of competing apple varieties. The simulated results imply that 91.60 percent of total consumer welfare changes are attributable to the increase in apple varieties. To be able to extrapolate our results to the entire U.S. apple market, we perform a back of the envelope analysis to extrapolate the estimates of welfare to the entire U.S. market and find that the introduction of Honeycrisp has increased total consumer welfare by about 940 million dollars during the study period. This corresponds to approximately 20 percent of the annual average domestic expenditures on public food and agricultural R&D.

The rest of this chapter is organized as follows. Section 3.2 gives a brief review of the literature. Section 3.3 describes the background of the U.S. apple market, followed by the data introduction in Section 3.4. We then present the analytical model and its underlying assumptions in Section 3.5. Section 3.6 discusses the identification strategy, as well as the estimation procedure. At last, we explain the results in Section 3.7 and conclude in Section 3.8.

3.2. Literature Review

Food and agricultural markets have become more consumer-oriented and consumer expectations of food quality are increasingly higher (McCluskey et al. 2007; Unnevehr et al. 2010). Higher awareness of healthy diets and changing consumer tastes, in turn, provides incentives for producers to improve the quality of their food and agricultural products. For example, Yue et al. (2013) surveyed grower preferences for fruit traits and find that growers prioritize quality traits, such as flavor, over horticultural traits, such as disease resistance. However, despite the increasing importance of demand-enhancing agricultural products, we do not have enough knowledge about their economic benefits. Unnevehr (1986) quantifies the changes in consumer welfare from improvements in the quality of rice and concludes that economic returns to agricultural research on grain quality are substantial. In another study, Brester et al. (1993) evaluate industry profits from the introduction of low-fat ground beef and find that the new product results in a small increase of less than 1 percent in equilibrium retail price and quantity of aggregate ground beef, as well as social welfare.

Estimation of demand systems is central in the research on measuring the economic benefits from the introduction of a new product.[15] Using the estimates from a demand model, some studies construct cost-of-living indices to summarize the total welfare changes resulting from a number new products. For example, Hausman (1999) investigates the bias of the Consumer Price Index (calculated by the Bureau of Labor Statistics) due to the omission of new products (i.e.,

[15] The literature on the demand estimation is large (e.g., Deaton and Muellbauer 1980; Hausman, Leonard, and Zona 1994; Berry 1994; Berry, Levinsohn, and Pakes 1995; Nevo 2000a, 2001). A review of these studies is beyond the scope of this paper. Nevo (2011) and Bonnet and Richards (2016) briefly survey the development of demand estimation.

cellular telephones). The author estimates the welfare changes due to the adoption of cellular telephones using a derived expenditure function from an estimated Hicksian demand. Similarly, Nevo (2003) develops a price index to account for the introduction of new products and quality changes in existing products based on the estimation of a brand-level demand system.

Other studies use estimated demand models to simulate market outcomes in certain counterfactual scenarios. For example, Hausman and Leonard (2002) estimate structural models of demand and supply to simulate the equilibrium prices in the absence of a new bath tissue product and then measure the difference in consumer welfare between the observed scenario and the counterfactual scenario. Similar approaches have been adopted to analyze the introduction of new products in a number of markets. Petrin (2002) evaluates welfare changes due to the introduction of minivans in the automobile market, Kim (2004) performs a similar analysis for the processed cheese market, and Pofahl and Richards (2009) quantify the consumer valuation of new products in the market for juice drinks. A notable distinction between these studies and our work is that we focus on a fresh produce item rather than a processed or highly industrial product. Although the private sector leads the breeding programs in many vegetable crops as well as some fruit crops, the development of new produce varieties is heavily influenced by leading research and germplasm innovation at the land-grant universities.[16] Therefore, our results inform policymakers and research institutions about insights into future public and private initiatives on agricultural investments.

This chapter is also related to the literature on consumer valuation of different apple varieties. Yue and Tong (2011) conduct a choice experiment in real markets with a follow-up survey to investigate the willingness to pay for new apple varieties versus existing ones. The authors find that there is a strong preference for new varieties and that new varieties with more

---

[16] In addition to apples, there are a large number of successful breeding programs in the land-grant universities, for example, strawberries at the University of California, Davis and the University of Florida, blueberries at the University of Florida, Michigan State University, and North Carolina State University, tomatoes at the University of Florida, and wheat at Kansas State University.

desired characteristics (e.g., firmness, crispness, and tartness) would receive higher premiums. Similarly, Rickard et al. (2013) develop an experiment to examine the impacts of names on a new apple variety. The results show that there is a price premium for using a sensory name for a new variety; however, changing the name of an existing variety has little influence. Other studies attempt to identify the internal quality characteristics that affect the consumer valuation of apples using individual surveys with contingent valuation questions. McCluskey et al. (2007) find consumers are willing to pay more for an apple with attributes closer to their subjective perceptions for texture, flavor, firmness, and tartness. In another study, McCluskey et al. (2013) measure the consumer valuation of internal quality characteristics across varieties and find that the willingness to pay for the same attribute is different by variety and associated with consumer demographics. A limitation of these studies is that findings are based on a small sample and the sample representativeness is questionable. Carew, Florkowski, and Smith (2012) evaluate the impacts of product characteristics on apple prices using a hedonic pricing model with Canadian sales data. The authors find that there is a price premium for a new apple variety and price premiums are positively correlated with the size and the grade of apples. Using market-level data, this chapter contributes to this line of research not only by providing new evidence on consumer valuation of apple characteristics and substitution patterns between the most popular varieties but also by evaluating the impacts of the introduction of a new apple variety on market shares and prices.

This chapter also contributes to the broad literature on returns to agricultural R&D. A review by Alston et al. (2009) indicates that a large number of studies have measured social returns to investments in agricultural R&D by identifying the lagged effect of research (the temporal attribution problem) and the spillover effect of new knowledge in certain areas (the spatial attribution problem). They find that returns to agricultural R&D primarily rely on the size of research-induced supply shifts and the scale of the affected industry. Therefore, prior studies typically estimate a supply function that enables them to measure the extent to which the supply curve shifts due to the agricultural R&D. Then, under certain assumptions, the welfare change from a downward shift of the supply curve against a stationary demand can be evaluated. Although the existing literature provides an abundance of evidence on the benefits of agricultural R&D to producers, there is a dearth of evidence on the benefits to consumers.

3.3. The Apple Market in the United States

Apples are one of the most popular fruits worldwide and apple varieties have been improved by cultivation and selection over thousands of years. Originally from Central Asia and widely grown in Asia and Europe, apples were brought to North America by early colonists, dating back to the 1630s. There are 7,500 varieties of apples grown around the world and 2,500 in the United States, of which more than 100 have been commercially sold at retail stores. According to Rickard (2013), an abundance of newly patented apples are under development and will be ready for introduction into the market. The records of the United States Patent and Trademark Office show that 156 patents of new apple varieties were approved during the period of 2000 to 2014.

As the second most valuable fruit on the market, the sales revenue of apples has exceeded two billion dollars since 2007 (USDA 2016b).[17] Apples are grown in all contiguous states but commercially produced in 32 states, led by Washington, New York, Michigan, Pennsylvania, California, and Virginia.[18] Most apples are sold fresh in retail stores. The sales quantity of fresh apples ranged from 6,300 to 7,900 million pounds between 2009 and 2014, about 70 percent of total production (USDA 2016b). After the adjustment for loss, the annual average consumption of fresh apples was 16.6 pounds per capita in 2014, up from 14.3 pounds in 2009 (USDA 2016a).

---

[17] By the Fruit and Tree Nut Yearbook (USDA 2016b), the four most-valuable fruits in the United States, are grapes, apples, oranges, and strawberries. The corresponding market sales in 2015 are, respectively, 5.56, 3.39, 2.22, and 1.96 billion dollars, which are summed up to 65 percent of total sales of fruits.

[18] Source: Apple Industry Statistics. United States Apple Association.

In contrast to processed food products sold by brand, apples are one of the few produce items marketed by variety. The sustainable growth of the apple industry is attributable to the development and commercialization of new varieties (Gallardo et al. 2012). In 2014, the top ten most purchased varieties accounted for 80 percent of total production.[19] Table 3.1 sketches the volume (pounds sold) market shares by variety in fall, the marketing/harvesting season of apples, from 2009 to 2014.[20] Gala is the most popular variety and accounts for one-third of total sales, while Honeycrisp is the fastest growing variety among the top five varieties. Table 3.2 shows that the annual average price of Honeycrisp has ranged between $1.85 and $2.30 per pound, which is about three times higher than the annual average price of Gala.

The Honeycrisp apple is a winter hardy variety developed by the apple breeding program at the University of Minnesota. After the 30-year breeding effort, it was introduced to the market in 1991 and rapidly became one of the most popular apples in the United States. In 2006, Honeycrisp was named the Minnesota State fruit. The patent protection of Honeycrisp in the United States expired in November 2008 and the University of Minnesota no longer earns royalties from the sales of Honeycrisp. But the sales in other countries where plant breeders' rights (similar to patent) and trademarks associated with Honeycrisp remain in force still generate a cash inflow to support future agricultural R&D programs at the university.

Honeycrisp apples are usually harvested in the early fall and sold until the early spring; they are not available in all seasons.[21] Figure 3.1 shows the annual sales quantity of Honeycrisp by season in the United States. The annual sales quantity increased fourfold from 36.02 million pounds

---

[19] The most-purchased apple varieties in 2014 were Gala, Red Delicious, Fuji, Granny Smith, Honeycrisp, Golden Delicious, McIntosh, Cripp's Pink/Pink Lady, Braeburn, and Jazz. Source: Retail Dietitian Toolkit. United States Apple Association.

[20] Due to seasonality in production, apple sales are significantly different by season.

[21] The sales season of Honeycrisp apples is usually from September to April. There is large variation in the sales of Honeycrisp across seasons, especially between summer and fall.

in 2009 to 127.62 million pounds in 2014. Meanwhile, the marketing season of Honeycrisp apples has been extended. The sales in spring began to rapidly increase in 2011, and the sales in summer had a jump between 2012 and 2013 although it was relatively small.

3.4. Data

The data come from several sources. The information on market prices and sales quantities are obtained from the retail point-of-sale scanner data collected by Information Resources, Inc. (IRI), known as IRI InfoScan Data.[22] The data contains weekly sales information of the representative retailers from 61 IRI cities across the United States in the period of 24 seasons from March 2009 to February 2015, where each IRI city is a collection of counties defined by the United States Census Bureau. Because of the seasonality in the apple market, we define the market as the combination of city and season. These IRI cities are denoted in Figure 3.2 by shadowed areas. The details of the data construction are given in Section 1 of Appendix B.

Given the information of product attributes based on the ingredient and nutrition labels, it is straightforward to define a characteristics space for most processed food products. However, it is not applicable for fresh apples, because the product quality and nutrient contents might vary with production factors, such as chemical usage, land quality, and weather condition. In fact, existing studies have shown that consumer valuations of apple varieties are dependent on texture and flavors, such as sweetness and tartness (e.g., McCluskey et al. 2007; Yue et al. 2013; McCluskey et al. 2013). Therefore, we project apple varieties onto a space characterized by attributes relevant to flavor and texture. The data including such attributes are obtained from the variety information provided by the Washington Apple Commission. The attribute data are only available for eight out of the top ten most-purchased apple varieties in the United States, including a continuous measure

[22] In particular, we use the primary IRI InfoScan Data, purchased by the United States Department of Agriculture and made available to academics for policy research.

of sweetness and a set of expert rankings for multiple uses of apples (e.g., pie stuffing, applesauce, baking, and freezing), which are used as proxy variables for texture.[23] Section 2 of Appendix B provides further details on apple characteristics.

A maintained assumption in the subsequent analysis is that apples are differentiated by variety and by retailer. In other words, the Gala apple sold by retailer A is considered as a different product from the Gala apple sold by retailer B. This assumption is plausible as it allows us to account for consumer heterogeneous preferences for retailer types. It, however, results in a dimensionality problem for parameters to be estimated due to a large number of differentiated apples in consumers' choice set. To address this problem, we first group the retailers by channel: convenience store, defense commissary store, dollar store, drug store, grocery store, and mass merchandise store. However, the data show that more than 85 percent of total apple sales occur in grocery stores. The data further show that grocery retailers significantly vary by size defined as the number of IRI cities in which a retailer owns a store. Therefore, we examine the distribution of the size of retailers and divide the retailers into four groups: local retailers, small regional retailers, regional retailers, and nationwide retailers. The details are discussed in Section 3 of Appendix B.[24]

To account for market expansion, the market share of an apple is defined by the division of its sales quantity over the total potential quantity in the market. Following previous literature (e.g., Nevo 2001; Kim 2004; Villas-Boas 2007), we assume the size of total potential quantity is proportional to the population in the IRI city with a cup of fruit per capita per day.[25, 26] Table 3.3

---

[23] The attribute data are available on the web page, http://bestapples.com/varieties-information/varieties/.

[24] Due to the privacy requirements of the data agreement with the Economic Research Service, USDA, we are unable to disclose the names of retailers in each category.

[25] Two cups of apple is equivalent to a large apple, which is about 0.5 pound. The relevant population is defined as the population covered by the data used in this chapter. Over the study period, the total quantities of apples sold by stores in our sample is about 13 percent of the total quantity sold by all retailers in the United States (the total apple sales is obtained from the USDA Economic Research Service). As a result, the proportionality factor for the population in the IRI city is 13 percent.

[26] Nevo (2001) assumes the market size is the total potential number of servings in a market where the potential is one serving of Ready-To-Eat breakfast cereal per capita per day. Kim (2004) calculates the market

presents the summary statistics for apple sales by retailer and by variety. Panel A of the table displays sample statistics for prices and market shares of apples. All nominal values are deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period in 1982-84. Panel B shows that consumers are more likely to buy fresh apples from local and small regional retailers than from regional and nationwide retailers. The market share of local retailers ranks first with an average of 13.29%, followed by small regional retailers with 11.59%, regional retailers with 4.44%, and nationwide retailers with 2.51%. Panel C provides the sample statistics of market shares by variety and shows that on average, Gala is the most popular variety and Honeycrisp is one of the top-five.

We obtain data on consumer demographics such as age and household income from the American Community Survey from 2009 to 2014 provided by the United States Census Bureau. To investigate if the younger generation is more likely to purchase new products than the older generation, we define a variable of young adult as a binary indicator for consumers aged between 25 and 44. In addition, we use retailer cost data as instruments for the estimation of demand. The cost information consists of apple prices by variety at the terminal markets and wage rates in the retailing industry. The price data from terminal markets are provided by the USDA Agricultural Marketing Service (AMS), including monthly average prices of different apple varieties paid by retailers in selected markets across the United States. There are 15 selected terminal markets across the United States and these markets are circled in Figure 3.2. Retailers in a city without a terminal market are assumed to pay the prices at the closest terminal market. The details of the construction of terminal market prices for every city are discussed in Section 4 of Appendix B. It is worth noting that Honeycrisp apples have the highest minimum price and widest price range in terminal markets.

---

size of processed cheese as a proportion to the size of the population with the proportional factor equal to one serving per capita per day. Villas-Boas (2007) defines the potential market of yogurt as half of the resident population in the market areas under the assumption that every individual consumes one half of a serving per week.

Table 3.4 presents summary statistics of cost-related variables. Wage rates for retailers in different cities are obtained from the BLS Occupational Employment Statistics Survey from 2009 to 2014. The survey reports wage rates at the state level for cashiers, truck drivers, tractor operators, stock movers, and packagers. Wage rates at the city level are averaged over states weighted by the associated population.[27]

3.5. Analytical Framework

3.5.1. Consumer Utility and Demand

We specify a discrete choice model of demand for apples (e.g., Berry 1994; Berry, Levinsohn, and Pakes 1995; Nevo 2001; Petrin 2002; Kim 2004). Let $j = 0, \dots, J$ denote differentiated apples, defined as a variety-retailer combination, with $j = 0$ indexing the outside option, $t = 1, \dots, T$ denote markets, defined as a city-season combination. The utility of consumer i from buying apple j in market t is

$$u_{ijt} = \boldsymbol{x}_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \boldsymbol{d}_t + \epsilon_{ijt}, \tag{16}$$

where p_{jt} is the price, $\boldsymbol{x}_{jt}$ is a $K \times 1$ vector of observed characteristics of apple j in market t, ξ_{jt} is the baseline utility of unobserved characteristics (i.e., unobserved valuation for econometricians but not consumers), $\boldsymbol{d}_t$ is a vector of dummies representing the seasonality in market t, and ϵ_{ijt} is an error term that is assumed to be independently and identically distributed (i.i.d.) across apples and be drawn from the Type I extreme value distribution. The conformable parameters (α_i, β_i) are the random coefficients to be estimated. These parameters represent consumer heterogeneous tastes for observed apple characteristics and prices, such that

[27] For example, the IRI city, Minneapolis-St. Paul (MSP), consists of counties in both Minnesota and Wisconsin. The wage rate of cashiers in MSP is hence averaged over the wage rates of cashiers in the two states by the associated population from MSP counties in Minnesota and Wisconsin.

$$\begin{pmatrix}\alpha_i\\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \Pi D_i + \Sigma v_i,$$

where the parameters (α, β) represent the homogenous tastes and the component $\Pi D_i + \Sigma v_i$ capture the individual discrepancies. For consumer i, the individual tastes are jointly determined by a $L \times 1$ vector D_i of demographic background variables (i.e., age and household income) and a $(1 + K) \times 1$ vector v_i of the idiosyncratic tastes, where Π and Σ are the corresponding parameter matrices with the dimension of $(1 + K) \times L$ and $(1 + K) \times (1 + K)$, respectively. To complete the demand model, the utility of consumer i from the outside option is specified as

$$u_{i0t} = \boldsymbol{d}_t + \epsilon_{i0t}.$$

The outside option includes other apples and fresh fruits sold in the stores included in this study, and any apple and fresh fruits sold in other stores. Following Nevo (2001), we denote θ_1 as a vector of the linear parameters (α, β) and θ_2 as a vector of nonlinear parameters $\left(vec(\Pi), vec(\Sigma)\right)$. The utility of consumer i can thus be written as

(17) $$u_{ijt} = \delta_{jt}\left(\boldsymbol{x}_{jt}, p_{jt}, \xi_{jt}, \boldsymbol{d}_t; \theta_1\right) + \mu_{ijt}\left(\boldsymbol{x}_{jt}, p_{jt}, D_i, v_i; \theta_2\right) + \epsilon_{ijt},$$

where δ_{jt} is the mean utility shared by all consumers, i.e., $\delta_{jt} = \boldsymbol{x}_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \boldsymbol{d}_t$, and μ_{ijt} is the consumer specific utility determined by the individual tastes, given as:

$$\begin{aligned}\mu_{ijt} &= \left[-p_{jt}, \boldsymbol{x}_{jt}\right](\Pi D_i + \Sigma v_i)\\ &= -p_{jt}\left(\pi_{p1} D_{i1} + \cdots + \pi_{pL} D_{iL} + \sigma_p v_{ip}\right)\\ &\quad + \sum_k x_{jt}^k (\pi_{k1} D_{i1} + \cdots + \pi_{kL} D_{iL} + \sigma_k v_{ik}).\end{aligned}$$

The heterogeneity is captured by the consumer specific utility, μ_{ijt}, as well as the idiosyncratic taste parameter, ϵ_{ijt}.

Let consumer i be characterized by a tuple $(D_i, v_i, \epsilon_{i\cdot t})$. The collection of consumers buying product j in market t is defined as a set C_{jt} such that

$$C_{jt} = \left\{ (D_{it}, v_{it}, \epsilon_{it}) \mid u_{ijt} \geq u_{ilt} \; \forall l = 0,1, \ldots, J \right\}.$$

Under the assumption that the distributions of demographics, idiosyncratic tastes, and error terms are independent, the market share of product j in market t is obtained as

$$s_{jt} = \int_{C_{jt}} dP(\epsilon)dP(v)dP(D), \tag{18}$$

where $P(\cdot)$ is the population distribution function. Note that if consumers are homogeneous in market t, i.e., $(D_{it}, v_{it}) = (\bar{D}_t, \bar{v}_t)$ and error terms are drawn from the Type I extreme value distribution, then (18) reduces to a classic (multinomial) logit model of demand.

3.5.2. Supply Side Model

To evaluate welfare changes due to the introduction of Honeycrisp apples, we must obtain equilibrium prices of other apples in a counterfactual scenario in which Honeycrisp apples would be absent in the market. To this end, we model the competition among retailers to derive equilibrium pricing rules. For estimation to be tractable we divide retailers into four groups based on their sizes.

Let J_r be a partition of apple varieties sold by a retailer group r. Given a vector $\boldsymbol{p}_{-r}$ of prices from rival groups, the retailer group r maximizes the group profit by jointly choosing a vector $\boldsymbol{p}_r$ of prices, that is,

$$\max_{\boldsymbol{p}_r} M \times \sum_{j \in J_r} (p_j - mc_j) s_j(\boldsymbol{p}_r, \boldsymbol{p}_{-r}),$$

where mc_j is the marginal cost of product j and M is the size of market. Suppose there exists a pure-strategy Bertrand-Nash equilibrium in prices. The optimal prices then satisfy the first order condition for apple $j \in \{1, \dots, J_r\}$,

$$s_j(\boldsymbol{p}_r) + \sum_{k \in J_r} (p_k - mc_k) \frac{\partial s_j(\boldsymbol{p}_r)}{\partial p_k} = 0,$$

which implies that the substitution patterns across apple varieties (i.e., own- and cross-price effects) are involved in the optimal pricing conditions. Let $\Delta^*(\boldsymbol{p})$ be defined as a matrix of substitution patterns such that $\Delta^*(\boldsymbol{p})_{jk} = -\partial s_j(\boldsymbol{p})/\,\partial p_k$, and Ω be defined as a matrix of ownership such that $\Omega_{jk} = 1$ if $j, k \in J_r$ and 0 otherwise. The first order conditions can be written in matrix notation,

$$\boldsymbol{s}(\boldsymbol{p}) - \Delta(\boldsymbol{p})(\boldsymbol{p} - \boldsymbol{mc}) = \boldsymbol{0}, \tag{19}$$

where $\boldsymbol{p}$ is a vector of prices, $\boldsymbol{mc}$ is a vector of marginal costs, $\boldsymbol{s}(\boldsymbol{p})$ is a vector of market shares, and $\Delta(\boldsymbol{p})$ is an element-wise product of ownership and the substitution matrix, i.e., $\Delta(\boldsymbol{p}) = \Omega * \Delta^*(\boldsymbol{p})$. Implied by (19), the vector of marginal costs is,

$$\boldsymbol{mc} = \boldsymbol{p} - \Delta(\boldsymbol{p})^{-1}\boldsymbol{s}(\boldsymbol{p}). \tag{20}$$

where the component $\Delta(\boldsymbol{p})^{-1}\boldsymbol{s}(\boldsymbol{p})$ captures the markup terms for the differentiated apples. Given the estimated demand model and observed prices, the marginal costs can be recovered by (20). Suppose that the marginal costs of apples are independent from the introduction of Honeycrisp. Then, the counterfactual prices can be obtained by using the first order conditions in (19) with the recovered marginal costs.[28]

Some non-trivial assumptions are made in the counterfactual analysis. First, the introduction of Honeycrisp apples would not affect the competition between retailer groups in the apple market. In other words, retailer groups are assumed to compete in Bertrand-Nash fashion regardless of the presence of Honeycrisp apples. Second, the demand model would not change in the counterfactual scenario. That is, we conduct the counterfactual analysis with the same demand estimates. It, however, does not imply that the substitution patterns are invariant. In fact, the substitution matrix $\Delta(\boldsymbol{p})$ is a function of market prices and hence vary with the equilibrium. Third,

[28] Retailers are allowed to adjust marginal costs systematically and proportionally to achieve economies of scale. That is, $\boldsymbol{mc}_1 = c \times \boldsymbol{mc}_0$ where $\boldsymbol{mc}_0$ and $\boldsymbol{mc}_1$ represent the marginal costs of other apples when the Honeycrisp was present and absent in the market respectively and c is a constant ratio implying the marginal costs of other apples would decrease due to the introduction of the Honeycrisp. In this chapter, we do not consider the economies of scale and set $c = 1$.

the value of the outside good is constant. It implies that the relative utility of an inside apple to the outside good would be the same if the attributes of the inside apple are fixed.

3.5.3. Evaluation of Consumer Welfare

Let $w > 0$ be a fixed expenditure on fresh fruits, $\boldsymbol{p}^{\text{with}}$ be a vector of prices when Honeycrisp apples are available in the market, and $\boldsymbol{p}^{\text{without}}$ be a vector of prices when absent. A consumer is strictly better off with the introduction of Honeycrisp apples if and only if

$$u(\boldsymbol{p}^{\text{with}}, w) - u(\boldsymbol{p}^{\text{without}}, w) > 0.$$

To measure the welfare changes in dollars, money metric indirect utility functions are employed. These functions are constructed using the means of consumer expenditures with fixed utility levels. A monetary measure of welfare changes for consumer i is

$$e\left(\overline{\boldsymbol{p}}, u_i(\boldsymbol{p}^{\text{with}}, w_i)\right) - e\left(\overline{\boldsymbol{p}}, u_i(\boldsymbol{p}^{\text{without}}, w_i)\right),$$

where $\overline{\boldsymbol{p}} \gg 0$ is an arbitrary price vector. Two natural choices for $\overline{\boldsymbol{p}}$ are price vectors $\boldsymbol{p}^{\text{with}}$ and $\boldsymbol{p}^{\text{without}}$. These two choices are equivalent under the assumption of no income effect. Following the literature (e.g., Nevo 2000b; Kim 2004), compensating variation (CV) is used to measure the welfare changes for consumer i,

$$CV_i = e\left(\boldsymbol{p}^{\text{with}}, u_i(\boldsymbol{p}^{\text{with}}, w_i)\right) - e\left(\boldsymbol{p}^{\text{with}}, u_i(\boldsymbol{p}^{\text{without}}, w_i)\right),$$

that is, $e\left(\boldsymbol{p}^{\text{with}}, u_i(\boldsymbol{p}^{\text{without}}, w_i)\right) = e\left(\boldsymbol{p}^{\text{with}}, u_i(\boldsymbol{p}^{\text{with}}, w_i)\right) - CV_i$, which implies

$$u_i(\boldsymbol{p}^{\text{without}}, w_i) = u_i(\boldsymbol{p}^{\text{with}}, w_i - CV_i).$$

Due to the linear specification of the indirect utility function, the compensating variation for consumer i can be written

$$CV_i = \frac{u_i(\boldsymbol{p}^{\text{with}}, w_i) - u_i(\boldsymbol{p}^{\text{without}}, w_i)}{\alpha_i},$$

where α_i is the constant marginal utility of income. Hence, the compensating variation for an average consumer can be calculated by

$$E[CV_i] = \int \frac{u_i^{\text{with}} - u_i^{\text{without}}}{\alpha_i} dP(\epsilon)dP(D)dP(v), \tag{21}$$

where u_i^{with} and u_i^{without} are the indirect utility functions with and without Honeycrisp apples and $u_i = \max_j u_{ij}$. With the assumption of the extreme value distribution for ϵ, McFadden (1981) provides the analytical solution to this integral,

$$E[CV_i] = \int \frac{\ln\left(\sum_{j=0}^{J^{\text{with}}} \exp\left(u_{ij}(\boldsymbol{p}^{\text{with}})\right)\right) - \ln\left(\sum_{j=0}^{J^{\text{without}}} \exp\left(u_{ij}(\boldsymbol{p}^{\text{without}})\right)\right)}{\alpha_i} dP(D)dP(v), \tag{22}$$

where $u_{ij}(p_j) = \boldsymbol{x}_j\beta_i - \alpha_i p_j + \xi_j$ is the utility level of consumer i from apple j evaluated at the price p_j.

The introduction of Honeycrisp apples affects the consumer welfare by increasing the number of apple varieties and changing prices for competing apples. To be able to measure these two impacts separately, the compensating variation can be decomposed as:

$$E[CV_i] = \int \left[\frac{\ln\left(\sum_{j=0}^{J^{\text{with}}} \exp\left(u_{ij}(\boldsymbol{p}^{\text{with}})\right)\right) - \ln\left(\sum_{j=0}^{J^{\text{without}}} \exp\left(u_{ij}(\boldsymbol{p}^{\text{with}})\right)\right)}{\alpha_i} + \frac{\ln\left(\sum_{j=0}^{J^{\text{without}}} \exp\left(u_{ij}(\boldsymbol{p}^{\text{with}})\right)\right) - \ln\left(\sum_{j=0}^{J^{\text{without}}} \exp\left(u_{ij}(\boldsymbol{p}^{\text{without}})\right)\right)}{\alpha_i}\right] dP(D)dP(v), \tag{23}$$

where the first term of the integrand represents the impact of the increase in apple varieties and the second term captures the impact of the change in prices of competing apples.

There is a caveat for the welfare analysis using the discrete choice model of demand with market level data. The welfare estimates might heavily rely on the idiosyncratic logit error due to the limited information of data (Petrin 2002). This problem arises from the assumption of the additive i.i.d. error in the random utility framework. It is clear in (23) that the direct impact of the introduction of Honeycrisp apples is always positive, since $J^{\text{with}} > J^{\text{without}}$ and $\exp(x) > 0$ for any x. In other words, consumers are always better off when Honeycrisp is in the apple market even if it is identical to a competing apple variety. As a result, the welfare impacts of the introduction of Honeycrisp apples could be overestimated. The random-coefficients model alleviates this problem, to a large extent, by separating consumers' heterogeneous tastes into two parts, ϵ_{ij} and μ_{ij}, where ϵ_{ij} is an individual error term and μ_{ij} is determined by the interactions between apple characteristics and consumer demographics. The compensating variation can be hence decomposed into the changes related to the error term ϵ_{ij} and the changes related to observed characteristics, $\delta_j + \mu_{ij}$.

3.6. Estimation

3.6.1. Endogenous Prices and Identification

A product price represents the implicit value of its characteristics, but not all characteristics are included in the demand estimation. As a consequence, the product prices are correlated with the estimation error through the consumer valuation of unobserved characteristics (also see Figure 3.3). This correlation raises the problem of price endogeneity, which is well-documented in prior studies (e.g., Berry 1994; Berry, Levinsohn, and Pakes 1995; Nevo 2000a, 2001). In this study, some taste characteristics of apples, such as crispness and juiciness, and appearance characteristics, such as size and color, are not included in the demand estimation, due to the limited variety information on apples. To see the impacts of unobserved variables on the demand estimation, consider a scenario wherein a consumer prefers only crispy apples. In other words, both apple prices and market demands are positively related to the crispness of apples. If crispiness is not controlled in the demand estimation, then it will be included in the error term. In turn, the positive correlation between the apple price and the error term biases downwards the estimate of price parameter α.

A regular remedy for the problem of endogenous prices is to use product-level instruments that are highly correlated with product prices but not correlated with unobserved characteristics. In order to find valid instruments, we need to understand the structure of product prices. Product prices are a function of marginal costs and a markup term, where the markup term represents the consumer valuation of all product characteristics. The variation in product prices can, thus, be divided into the exogenous variation in marginal costs and the endogenous variation in consumer valuation. Therefore, valid instruments are required to identify price variation through the changes in marginal costs.

There are three sets of cost-related variables employed in the estimation. First, by exploiting the panel structure of the data, Hausman (1994) and Nevo (2000a, 2001) calculate the average product prices in all other cities to capture the changes in marginal costs. These average product prices are viable instruments under the assumption that cross-city demand shocks (i.e., the change in unobserved valuation) are independent across cities. This assumption, however, is tenuous if there are cross-city advertising and promotion activities. To accommodate the potential problem of related marketing strategies across cities within a Census-defined division, we replace the average product prices over all other cities with the average product prices over cities in all other divisions.[29] Although these instruments by construction are not affected by cross-city unobservables, their exogeneity would still be questionable if there exist systematic demand shocks. For example, a sudden awareness of some nutrients in apples would increase the unobserved valuation of apples and hence the market demand across the United States. However, these demand shocks are not much of a concern in our case because all the apple varieties included in our analysis

[29] The United States Census Bureau defines four statistical regions with nine divisions for data collection and analysis. Perez et al. (2001) find that the patterns of apple consumptions are different across Census-defined regions.

are well-established.[30] In addition, we use period dummies in the demand model that capture any time-variant national shocks (Hausman and Leonard 2002).

Second, we use terminal market prices of different apple varieties to represent the retailer costs for apples. Following Villas-Boas (2007), we interact these prices with retailer group dummies and hence obtain product-level instruments. These instruments capture the differences in costs of apples and account for the variation in prices due to the changes in marginal costs by the combination of variety and retailer.

Third, we use wage rates in the retailing industry as another set of instruments. Retailers are assumed to be price takers in the labor market. These pre-determined costs of variable inputs would influence retailers' marketing strategies (including pricing conditions) but not consumer valuation for unobserved characteristics of apples. Thus, cross-city wage rates of labor, such as cashier, truck drivers, tractor operators, stock movers, truck loaders, and packagers, are viable instruments to disentangle the cross-city variation in marginal costs from the variation in the consumer valuation.

3.6.2. Demand Estimation

The demand model is estimated using the generalized method of moments (GMM) and the estimates of parameters are determined to minimize the differences between the observed and the predicted market shares of apples. Calculating the integral in (3) raises a challenge for applying instruments to the endogenous apple prices, which are correlated with the consumer valuation of unobserved characteristics ξ_{jt}. The key to this challenge is to recover the mean utility δ_{jt} and construct the moments (i.e., orthogonal conditions) for ξ_{jt}. Berry (1994) provides an inversion

[30] To the best of our knowledge, there has not been nationwide news on apple nutrients reported over the study period.

method to obtain δ_{jt} in the (multinomial) logit model by matching the observed market share s_{jt}^{obs} with the predicted market share $s_{jt} = \delta_{jt}/\left(1 + \sum_{k=1}^{J} \delta_{kt}\right)$. The solution to δ_{jt} is of an analytical form such that $\delta_{jt} = \log\left(s_{jt}^{\text{obs}}\right) - \log\left(s_{0t}^{\text{obs}}\right)$. However, this analytical inversion method is impeded by the integral in the random-coefficients logit model. As a result, a numerical inversion method developed by Berry et al. (1995) is employed and the value of δ_{jt} depends on the non-linear parameters, θ_2. Suppose X_{jt} is a matrix of variables contained in the mean utility. Then the linear parameters, θ_1 can be expressed as $\theta_1 = \left(X_{jt}' X_{jt}\right)^{-1} X_{jt}' \delta_{jt}(\theta_2)$, which suggests that θ_1 is a function of θ_2. Let Z be a $n \times L$ matrix of instruments and $\xi(\theta_2)$ be a $n \times 1$ vector of the consumer valuation of unobserved characteristics. The estimation is, therefore, to find optimal θ_2^* such that

$$\theta_2^* = \arg\min_{\theta_2} \{G' W^{-1} G\}$$

where G is a sample moment $G(\theta_2) = (1/n) Z' \xi(\theta_2)$ and W is a consistent estimate of the asymptotic variance of $\sqrt{n} G(\theta_2)$. The estimation follows Nevo's (2000a, 2001) procedure using a simulated GMM objective function with analytical gradients.

3.7. Results

3.7.1. Parameter Estimates and Elasticities

We first estimate a logit model of demand for apples to explore viable specifications for the full model (i.e., random-coefficients logit model) and illustrate the problem of endogenous prices and the need for instruments. The estimates of the logit model are presented in Table 3.5, where the dependent variable is given by $\log\left(s_{jt}\right) - \log(s_{0t})$. The OLS results in columns 1 and 2 show that there is a small difference between estimates of price coefficients. This suggests the city-specific variables of average consumer demographics are significant but provide little information on the cross-product variation in mean utilities.[31] Columns 3 to 8 show that the orthogonality conditions for product prices might be violated and the specifications with instruments are preferable to the simple regression. Compared to the OLS results, the IV estimates of the price parameter are substantially larger in absolute value. This implies that product prices are negatively associated with mean utility but positively (negatively) with the consumer valuation of favorable (unfavorable) unobserved characteristics. The inclusion of endogenous prices without instruments leads to a relatively inelastic demand for apples. Compared to the estimates in column 3 (or 6), the estimates of the price parameter in columns 4 and 5 (or 7 and 8) are smaller in absolute value when the average product prices outside the division are used as instruments. The retailer costs, measured by terminal market prices and wage rates in the retailing industry, are included to account for the cross-variety variation in prices and the cross-city variation in prices. We also find that estimates of the price parameter are robust to adding city-average consumer demographics. In addition, the adjusted R-squared and the F-statistic for the exclusion of instruments in the first stage regression suggest the weak instruments are less of a concern.

Table 3.6 displays estimates from the full model (i.e., the random-coefficients logit model of demand) based on (17) with different specifications. Consumer heterogeneity is characterized by consumer demographics and idiosyncratic shocks.[32] The inclusion of demographic variables creates a scaling problem because of differences in units. To address this issue, we apply the logarithm transformation to age and household income and express all demographic variables as the deviations from the mean (e.g., Nevo 2000a; Kim 2004; Villas-Boas 2007). The full model is

---

[31] The term “product” hereafter refers to the differentiated apple.

[32] In every market (i.e., the combination of city and season), we simulate 1,000 consumers characterized by age, household income, and idiosyncratic tastes. The demographic variables of age and household income are sampled from the associated empirical distributions in the American Community Surveys. In line with Petrin (2002), we draw consumers’ idiosyncratic tastes from the normal distribution truncated at 95 percent based on two reasons: these tastes are bounded above and below, and the distributions of consumers’ preferences on apple characteristics are balanced.

estimated with all sets of instruments, including average product prices outside the division, terminal market prices, and wage rates in the retailing industry.

The linear parameters, θ_1, are the mean of random coefficients estimated by the minimum-distance procedure (e.g., Nevo 2000a, 2001), wherein the product-fixed effects are regressed on apple characteristics. The estimates of linear parameters are statistically significant with expected signs and are robust to alternative demand specifications. These negative price coefficients are greater in absolute value than those estimates from the logit model with the same set of instruments, implying more elastic demand for apples. The coefficients of observed apple characteristics suggest that consumers prefer the varieties well-suited for making applesauce and baking over the varieties well-suited for freezing and having a high degree of sweetness. In addition, the coefficients of retailer group reveal that consumers are more likely to buy apples from a small regional retailer than a nationwide retailer.

The nonlinear parameters, θ_2, capture the heterogeneity in consumer preferences and tastes. The deviations from the homogenous tastes for apple characteristics are allowed to vary with demographic variables and idiosyncratic shocks. The coefficients of idiosyncratic shocks, age, and young adult, however, are not statistically significant. This implies that the heterogeneity in consumer preferences and tastes might not arise from the idiosyncratic shocks and the apple consumption patterns are not remarkably different by age and generation.[33] On the contrary, the coefficients of the interaction terms between apple characteristics and household income are statistically significant. This suggests that the variation in consumer preferences for apple characteristics is primarily determined by the variation in household income. In particular, the positive coefficient of the interaction term between household income and the small regional retailer group implies that the marginal utility of shopping from a small regional retailer increases

---

[33] Additional specifications that are not presented in this chapter also show that the coefficient of age-related variables are not statistically significant in the demand model.

with household income. The suitability for freezing is not a favorable apple characteristic for consumers with above average household income. Besides, there is a quadratic impact of household income on consumer disutility of apple prices. The coefficient signs of the interaction terms between price and household income and household income-squared are opposite. This implies that consumer insensitivity for apple prices is increasing in household income at a diminishing rate.

Next, we discuss elasticity estimates and substitution patterns between apple varieties. The own- and cross-price elasticities are associated with the empirical distribution of consumer demographics, product fixed effects, and seasonality. Because of the large dimension, we only present the summary statistics of the estimates of elasticities by variety and by retailer group in Table 3.7. The results show that the estimated demand curves for apple varieties are highly elastic with respect to own price. On average, the most-purchased variety, Gala apples, has the least elastic demand, which is about half of the own-price elasticity of Honeycrisp apples. Specifically, a one percent decrease in the own price will increase the sales quantity of Gala apples by 2.80 percent and Honeycrisp apples by 5.58 percent. Moreover, the own-price elasticities vary across retailer groups. On average, the group of nationwide retailers has lower own-price elasticities than other groups.

Table 3.7 also shows that cross-price elasticities of apples from the same retailer group are generally greater than those from a different group. In line with Kim (2004) and Villas-Boas (2007), we find that cross-price elasticities are relatively smaller than own-price elasticities. For example, the sales quantity of Golden Delicious will increase by 0.09 percent if the average price over other apples from the same retailer group increase by one percent, while it will increase by 0.06 percent if the average price over other apples from different retailer groups increases by one percent. It implies that consumers are more likely to substitute one variety for another in the same retailer group than in a different one. The only exceptions are consumers who buy apples from nationwide retailers. In addition, the large standard deviations of cross-price elasticities imply that cross-price elasticities would change within a wide range and the specification of the logit model is too restrictive (Villas-Boas 2007).

3.7.2. Counterfactual Analysis

Using the retailers' pricing rules and the estimated demand elasticities, we simulate the market outcomes in a counterfactual scenario in which Honeycrisp apples are removed from the market. The analysis focuses on markets where market shares of Honeycrisp apples are greater than or equal to 1 percent. The price changes of competing apple varieties due to the introduction of Honeycrisp are presented in Table 3.8, where these prices are averaged across retailer groups, weighted by sales quantity. The upper half of Table 3.8 displays the average price changes by variety in 481 markets. The prices of apples in most markets decline in response to the introduction of Honeycrisp. For example, the average price of Gala decreases by 0.72 percent, or a drop of 0.27 cents per pound. In contrast to Gala, Golden Delicious exhibits the least responsiveness to the introduction of Honeycrisp. The last column shows the number of markets where the introduction of Honeycrisp increases the competition and reduces the prices of competing apple varieties. Moreover, the impact of the introduction of Honeycrisp on the decline in prices of competing apple varieties is positively correlated with the market share of Honeycrisp. The lower panel of Table 3.8 presents the equilibrium prices of competing apple varieties in 96 markets where market shares of Honeycrisp are greater than or equal to 5 percent. In this situation, the average price of Gala decreases by 2.23 percent, or a drop of 0.71 cents per pound.

Next, we calculate changes in the market shares of competing apple varieties, the overall market size, and the total sales revenue. The estimates in Table 3.9 are based on the sample of 481 markets. The results show that Honeycrisp has increased both the total sales quantity and the total sales revenue. In the study period, the number of markets with market shares of Honeycrisp greater than or equal to 1 percent increased from 42 in 2009 to 111 in 2014, and the total sales quantity of

Honeycrisp increased from 13.35 million pounds in 2009 to 47.66 in 2014. Compared to the counterfactual results, the introduction of Honeycrisp leads to a decrease in the total sales of other apples but an overall increase in the total sales of all apples. Table 3.9 shows that the total sales of other apples was 57.40 million pounds less than the total sales of other apples when Honeycrisp is removed from the markets (i.e., Counterfactual Total Quantity minus Other Apples' Quantity), and that the introduction of Honeycrisp increased the total sales of all apples by 8.03 percent from 1,574.90 to 1,701.45 million pounds. These results suggest that Honeycrisp attracted more consumers who would otherwise choose the outside option. Besides, the gain in sales revenue due to the introduction of Honeycrisp outweighed the loss in sales revenue due to the decline in the prices of others. As the total sales quantity rose in the study period, the total sales revenue increased by 21.25 percent from 621.65 to 753.76 million dollars.

We evaluate the changes in consumer welfare using the measure of CV. The CV suggests the pecuniary change for consumers so that they are indifferent between the observed scenario (i.e., the data) and the counterfactual scenario. In other words, the CV measures the amount of money a Honeycrisp consumer needs to be compensated in the counterfactual scenario to maintain the same utility as before (i.e., the utility achieved when Honeycrisp apples are in the market). Before delving into the CV measure for consumer welfare, we examine the extent to which the assumption of additive i.i.d. error in the random utility framework affects the estimates of welfare change. Table 3.10 shows the decomposition of welfare changes for an average consumer into a component from observed characteristics, $\delta_j + \mu_{ij}$, and a component from the logit error, ϵ_{ij}. For the markets with market shares of Honeycrisp greater than or equal to 1 percent, the total average changes in consumer welfare due to the introduction of Honeycrisp are 3.14 cents per pound, of which 59.55 percent (i.e., 1.87 cents per pound) are related to the changes from observed characteristics and 40.45 percent (i.e., 1.27 cents per pound) to the changes from the logit error. In addition, the results show that the percentage change in consumer welfare stemming from the observed characteristics is positively associated with the market share of Honeycrisp. For the markets with Honeycrisp share greater than or equal to 5 percent, 70.82 percent of the total average changes in consumer welfare can be explained by the changes from observed characteristics. Thereby, the problem due to the assumption of additive i.i.d. error is less of a concern in this study.[34] The total welfare change due to the introduction of Honeycrisp is calculated by

$$\text{Total Change in Consumer Welfare} = \sum_t E[CV_{it}] \times Q_t$$

where Q_t is the total sales quantity of apples in market t. Table 3.11 shows that the total benefits in consumer welfare increased from 3.03 million dollars in 2009 to 15.20 in 2014. The decomposition of total changes in consumer welfare suggests that the growth of consumer welfare is primarily attributable to the increase in apple varieties rather than price competition. The total consumer welfare gains from Honeycrisp increased from 2.76 million dollars in 2009 to 13.91 million dollars in 2014, corroborating the recent growth of the Honeycrisp demand and popularity in the United States.

3.8. Conclusion

Agricultural research and development programs on new demand-enhancing products have become increasingly important over the past decade. Large numbers of new agricultural products have been developed and introduced in the United States to serve consumers' increasing expectations of food quality. However, little is known about their economic benefits. In this paper, we analyze the welfare impacts of the introduction of Honeycrisp apples using structural models of consumer demand and retailer supply.

---

[34] Petrin (2002) finds the welfare analysis can be improved by augmenting the full model with additional micro-level data of households.

We estimate consumer demand for apples in a discrete choice approach with random coefficients, and model the retailer competition in the Bertrand-Nash setting. The model addresses the problem of endogenous product prices and explicitly incorporates consumer heterogeneous tastes. With both demand estimates and retailers' pricing rules, we predict counterfactual prices of competing apple varieties in the absence of Honeycrisp and evaluate the changes in consumer welfare and total sales quantity and revenue.

The main results show that consumers are better off in a market with more options of apple varieties. For the sample markets, we find that the introduction of Honeycrisp has increased consumer welfare by 3.14 cents per pound on average, corresponding to a total of 49.03 million dollars overall the study period. More than 90 percent of welfare gain is explained by the increased number of total apple varieties, while the remaining part is explained by the decline in prices of competing apple varieties. The extent of the decline is positively associated with the market share of Honeycrisp. We also find that the introduction of Honeycrisp has increased the total sales of all apples by 126.48 million pounds and the total sales revenue by 132.12 million dollars, which are equivalent to 8.03 percent and 21.25 percent of their corresponding counterfactual estimates, respectively.

It is important to put the magnitude of the estimated welfare change into context. Suppose the estimated welfare change from our sample can be extrapolated to the entire U.S. apple market. In that case, a back of the envelope analysis suggests the introduction of Honeycrisp has increased total consumer welfare in the United States by about 940 million dollars between 2009 and 2014.[35] This gain corresponds to 21 percent of the annual average domestic expenditures on public food

[35] The estimated change in total consumer welfare is a product of total apple sales quantity and the change in average consumer welfare of buying apples (i.e., $E[CV_i]$). According to the USDA Food Availability (Per Capita) Data System, the number of total sales quantity in the U.S. apple market between 2009 and 2014 was 29,933.09 million pounds. Given that $E[CV_i]$ is 3.14 cents per pound, the estimated increase in total consumer welfare was 939.90 million dollars.

and agricultural R&D between 2000 and 2011 in the United States (Pardey et al. 2016).[36] Aligned with previous literature, our estimates also imply that there are substantially large returns to agricultural R&D.[37]

Due to the lack of disaggregated data on apple production, we do not investigate the vertical relationship between retailers and growers on the supply side. As a result, the model only accounts for the welfare changes in the total sales revenue of retailers rather than growers. The price premium paid for Honeycrisp strongly motivates growers to produce more Honeycrisp apples. The estimated increase in apple sales revenue is consistent with recent growers' planting reports. In addition, news articles in New York Times and on National Public Radio claim that many growers in Washington state have been looking to switch from Gala and Red Delicious to Cosmic Crisp, a new variety derived from Honeycrisp (Karp 2015; Charles 2017). This is in line with our finding that Gala and Red Delicious are the two varieties that suffer the largest decreases in prices from the introduction of Honeycrisp. Nevertheless, the incentives might quickly vanish as the growth of the Honeycrisp production will eventually reduce its price premium.

---

[36] Pardey et al. (2016) report that the average annual domestic expenditures on public food and agricultural R&D is about 4.47 billion dollars in United States between 2000 and 2011.

[37] As Bresnahan points out relating to Hausman's (1996) study on the valuation of new goods, the large consumer surplus results from a steep demand for the new variety (i.e., Honeycrisp) and its small substitutability between other varieties.

Table 3.1. Apple Market Shares by Variety (Percent of Total Volume)

Variety	2009-Fall	2010-Fall	2011-Fall	2012-Fall	2013-Fall	2014-Fall
Gala	27.58	30.50	32.89	30.58	30.30	31.46
Red Delicious	21.66	21.66	18.18	19.11	15.73	13.25
Fuji	9.57	8.02	8.82	9.98	10.26	11.51
Granny Smith	10.70	10.48	10.82	10.69	9.50	10.75
Honeycrisp	3.81	5.83	6.63	6.34	6.79	8.56
Golden Delicious	5.74	4.84	4.32	3.78	3.58	3.48
Mcintosh	6.03	5.52	5.25	4.38	4.81	4.82
Pink Lady/Cripps Pink	0.51	0.56	0.45	1.03	1.51	1.05
Braeburn	1.19	1.22	0.61	0.80	0.71	0.67
Jazz/Scifresh	0.35	0.61	0.36	0.88	0.98	1.10
Top 5	73.31	76.50	77.33	76.70	72.59	75.53
Top 10	87.14	89.24	88.31	87.58	84.17	86.65

Source: IRI Infoscan Data.

Table 3.2. Apple Market Prices by Variety (Dollars per Pound)

Variety	2009	2010	2011	2012	2013	2014
Gala	0.63	0.60	0.63	0.70	0.70	0.65
Red Delicious	0.50	0.52	0.56	0.58	0.61	0.64
Fuji	0.68	0.76	0.76	0.82	0.76	0.83
Granny Smith	0.77	0.80	0.80	0.89	0.89	0.85
Honeycrisp	2.11	1.85	1.97	2.30	2.24	2.07
Golden Delicious	0.89	0.91	0.98	1.06	1.00	0.90
Mcintosh	0.63	0.63	0.68	0.79	0.67	0.62
Pink Lady/Cripps Pink	1.26	1.28	1.20	1.25	1.18	1.17
Braeburn	1.16	1.23	1.29	1.46	1.51	1.55
Jazz/Scifresh	1.75	1.46	1.17	1.10	1.24	1.21

Source: IRI Infoscan Data.

Table 3.3. Summary Statistics of Apple Sales by Retailer Group and Variety

	Mean	Median	SD	Min	Max
Panel A. Price (Dollars per Pound) and Market Share (Percent)					
Price	0.54	0.47	0.30	0.05	2.04
Market Share [a]	1.07	0.31	2.24	0.00	48.23
Panel B. Market Shares by Retailer (Percent)					
Local	13.29	9.18	13.59	0.00	86.48
Small Regional	11.59	3.77	13.72	0.00	50.92
Regional	4.44	2.82	5.65	0.00	35.07
Nationwide	2.51	1.78	3.22	0.00	29.44
Panel C. Market Shares by Variety (Percent)					
Gala	5.45	2.92	6.37	0.00	48.48
Red Delicious	4.56	2.92	4.60	0.00	44.35
Fuji	2.21	1.30	2.94	0.00	20.40
Granny Smith	3.06	1.88	3.03	0.00	15.35
Honeycrisp	1.45	0.46	2.73	0.00	24.10
Golden Delicious	1.39	0.77	1.72	0.00	14.62
Pink Lady/Cripps Pink	0.79	0.46	1.05	0.00	8.63
Braeburn	0.69	0.28	1.52	0.00	20.37

Note: Prices are deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period at 1982-84.

[a] Market share is defined as the ratio of the apple sales quantity to the potential market size. The potential market size is defined in footnote 12.

Table 3.4. Summary Statistics for Cost Information

	Mean	Median	SD	Min	Max
Apple Prices in the Terminal Markets (Dollars per Pound)					
Minimum Prices					
Braeburn	0.28	0.27	0.06	0.10	0.53
Fuji	0.23	0.23	0.05	0.10	0.41
Gala	0.25	0.24	0.05	0.12	0.40
Golden Delicious	0.21	0.21	0.05	0.11	0.36
Granny Smith	0.25	0.25	0.05	0.11	0.40
Honeycrisp	0.43	0.39	0.19	0.12	1.60
Pink Lady/Cripps Pink	0.32	0.31	0.08	0.11	0.58
Red Delicious	0.21	0.20	0.05	0.08	0.34
Range of Prices					
Braeburn	0.15	0.12	0.11	0.00	0.63
Fuji	0.23	0.20	0.12	0.01	0.72
Gala	0.22	0.20	0.11	0.00	0.74
Golden Delicious	0.18	0.17	0.07	0.04	0.44
Granny Smith	0.20	0.17	0.12	0.03	0.78
Honeycrisp	0.25	0.21	0.24	0.00	1.23
Pink Lady/Cripps Pink	0.18	0.16	0.13	0.00	0.70
Red Delicious	0.16	0.14	0.08	0.04	0.62
Relevant Labor Costs in the Retailing Industry (Dollars per Hour)					
Minimum Wage Rates					
Cashiers	3.48	3.49	0.20	3.07	4.07
Heavy Truck Drivers	5.48	5.45	0.44	4.29	7.22
Light Truck Drivers	3.85	3.84	0.22	3.36	4.48
Tractor Operators	4.45	4.45	0.29	3.62	5.38
Stock Movers	3.68	3.68	0.17	3.30	4.13
Packagers	3.53	3.53	0.18	3.17	4.07
Range of Wage Rates					
Cashiers	2.20	1.91	0.73	1.39	4.67
Heavy Truck Drivers	7.07	7.06	0.73	4.57	9.30
Light Truck Drivers	8.32	8.37	0.89	5.40	10.33
Tractor Operators	5.41	5.21	0.96	3.64	8.80
Stock Movers	4.71	4.68	0.54	3.35	6.45
Packagers	3.46	3.54	0.56	2.01	4.94

Note: Prices and wage rates are deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period at 1982-84. Apple prices by variety in the terminal markets are provided by the USDA Agricultural Marketing Service. Minimum prices are defined as the 5th percentile price and the ranges are defined as the associated differences between the 5th and the 95th percentile price. Relevant labor costs in the retailing industry are obtained from the BLS Occupational Employment Statistics Survey. Minimum wage rates are defined as the 10th percentile wage rate and the ranges are defined as the associated differences between the 10th and the 90th percentile wage rate.

Table 3.5. Results from the Logit Model

	OLS		IV					
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p_{jt}	–2.075***	–2.127***	–12.882***	–7.320***	–7.324***	–11.831***	–7.358***	–7.305***
	(0.105)	(0.104)	(1.185)	(0.255)	(0.246)	(1.055)	(0.252)	(0.241)
Mean of Young Adult Ratio		–4.987***				–3.322*	–3.488*	–3.595**
		(1.922)				(1.987)	(1.952)	(1.770)
Mean of log(Age)		–4.434**				–9.316***	–4.964***	–6.819***
		(1.826)				(2.040)	(1.867)	(1.737)
Mean of log(Income)		50.326***				48.853***	55.185***	54.397***
		(10.069)				(11.306)	(10.534)	(10.167)
Mean of log(Income)^2		–2.160***				–2.094***	–2.365***	–2.336***
		(0.459)				(0.515)	(0.479)	(0.463)
Fit/Exogeneity Test for p_{jt}	0.250	0.274	89.777	271.325	216.306	98.636	270.079	225.499
Frist Stage Regression								
Adjusted R^2			0.596	0.643	0.648	0.599	0.645	0.651
F-statistic for Instruments			14.782	94.057	63.003	15.722	94.572	63.608
Average Prices Outside Division			No	Yes	Yes	No	Yes	Yes
Terminal Market Prices			Yes	No	Yes	Yes	No	Yes
Wage Rates			Yes	No	Yes	Yes	No	Yes

Note: The sample size is 26,089. The dependent variable is given by $\log(s_{jt}) - \log(s_{0t})$. All specifications include the product fixed effects and the period dummies. The null hypothesis of exogeneity test is that p_{jt} is exogenous. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table 3.6. Results from the Full Model

	Variable	(1)	(2)	(3)	(4)
Mean	Price	–11.048(1.800)***	–11.309(2.169)***	–11.052(2.782)***	–11.045(3.112)***
	Constant	–3.123(0.176)***	–2.480(0.275)***	–2.410(0.288)***	–2.011(0.281)***
	Sauce	0.525(0.065)***	0.228(0.041)***	0.477(0.105)***	0.432(0.107)***
	Baking	2.446(0.173)***	1.493(0.099)***	3.296(0.434)***	4.287(0.446)***
	Freezing	–5.982(0.704)***	–4.324(0.720)***	–6.070(1.047)***	–7.441(1.045)***
	Sweetness	–2.499(0.399)***	–2.393(0.366)***	–2.930(0.526)***	–4.642(0.525)***
	Local	3.204(0.099)***	3.102(0.121)***	3.011(0.164)***	3.177(0.165)***
	Small	4.298(0.348)***	3.961(0.335)***	3.942(0.550)***	6.732(0.638)***
	Regional	2.319(0.095)***	0.781(0.848)	2.863(0.711)***	3.733(0.714)***
Interaction w.	Price		0.072(9.045)	0.076(11.534)	0.075(13.081)
Shocks	Constant		–0.086(6.940)	–0.091(5.962)	–0.091(6.232)
	Sauce			–0.032(6.183)	–0.031(7.205)
	Baking			–0.064(13.146)	–0.065(15.965)
	Freezing		–0.087(8.462)	–0.048(9.956)	–0.048(13.090)
	Sweetness		–0.067(12.414)	–0.041(7.732)	–0.042(9.291)
	Local		–0.020(20.600)	0.018(10.345)	0.019(12.567)
	Small		–0.010(20.828)	0.041(41.967)	0.041(53.321)
	Regional		0.138(17.905)	0.118(13.765)	0.119(14.518)
Interaction w.	Price	0.055(17.856)	0.071(23.635)	0.051(26.435)	–0.091(44.344)
Young Adult	Constant				0.123(30.790)
Interaction w.	Price	0.001(15.654)	0.007(19.832)	–0.0004(26.361)	–0.209(54.739)
Age	Constant				0.178(35.595)
Interaction w.	Price	151.714(26.564)***	148.008(40.987)***	152.080(46.764)***	152.128(37.146)***
Income	Constant	4.067(5.478)	3.913(5.133)	4.071(8.582)	4.085(9.722)

	Sauce	1.287(4.534)		1.269(5.843)	1.263(6.314)
	Baking	4.729(4.636)		4.765(5.202)	4.783(5.930)
	Freezing	–20.761(7.100)***	–14.928(6.699)**	–20.843(10.353)**	–20.879(11.841)*
	Sweetness	–12.587(6.819)*	–9.316(7.001)	–12.616(10.178)	–12.658(11.037)
	Local	2.306(2.991)	0.836(3.031)	2.327(4.439)	2.336(4.929)
	Small	10.929(3.801)***	10.258(4.593)**	10.985(5.210)**	10.994(5.399)**
	Regional	2.085(3.478)	1.994(5.596)	2.143(6.595)	2.161(6.914)
Inter. w. Inc^2	Price	–7.333(1.296)***	–7.139(2.004)***	–7.350(2.245)***	–7.352(1.797)***
GMM		941.826	979.350	939.538	939.265
R^2 Min.		0.898	0.821	0.806	0.818
Price Coef. > 0		0%	0%	0%	0%

Note: The sample size is 26,089. All specifications include the period dummies and use the same set of instruments (i.e., the average prices outside the division overall seasons, the terminal market prices, and the relevant wage rates in the retailing industry). The parameters of apple characteristics are estimated by the minimum-distance procedure. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table 3.7. Estimates of Own- and Cross-Price Elasticities

	Own-Price		Cross-Price					
			Same Retailer Group (Yes)		Same Retailer Group (No)		Average	
Variety	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Braeburn	–3.958	2.191	0.040	0.107	0.026	0.060	0.031	0.081
Fuji	–3.855	1.138	0.052	0.104	0.053	0.162	0.053	0.144
Gala	–2.802	1.004	0.050	0.112	0.043	0.135	0.045	0.128
Golden Delicious	–6.466	2.943	0.093	0.272	0.061	0.173	0.072	0.214
Granny Smith	–3.737	1.608	0.063	0.155	0.053	0.163	0.056	0.160
Honeycrisp	–5.584	2.712	0.030	0.086	0.021	0.052	0.024	0.067
Pink Lady/Cripps Pink	–5.639	2.972	0.050	0.192	0.027	0.081	0.035	0.133
Red Delicious	–2.961	2.011	0.033	0.135	0.022	0.060	0.026	0.094
Retailer Group								
Local	–4.243	1.998	0.072	0.193	0.026	0.079	0.043	0.136
Small regional	–5.925	3.213	0.076	0.215	0.023	0.076	0.041	0.140
Regional	–4.560	2.134	0.036	0.102	0.030	0.086	0.032	0.092
Nationwide	–3.377	2.077	0.026	0.070	0.064	0.180	0.050	0.149

Note: Means and standard deviation of estimated own- and cross-price elasticities are presented here. The third and the fourth column show the statistics for estimated elasticities only from the same retailer group by variety and by the type of retailer group. The fifth and the sixth column show the similar information but from different retailer groups. The last two columns show the overall average across retailer groups.

Table 3.8. Equilibrium Prices (Cent per Pound) with and without Honeycrisp

	Price	C. Price	Price Change	Number of Markets where Price ≤ C. Price (Percent in Total)
Market Shares of the Honeycrisp ≥ 1 percent (481 Markets)				
Braeburn	67.86	67.98	0.12 (0.18%)	351 (73%)
Fuji	54.05	54.17	0.13 (0.22%)	389 (81%)
Gala	37.28	37.55	0.27 (0.72%)	473 (98%)
Golden Delicious	54.16	54.18	0.02 (0.04%)	291 (61%)
Granny Smith	41.90	41.96	0.06 (0.14%)	317 (66%)
Pink Lady/Cripps Pink	63.84	64.01	0.17 (0.27%)	370 (77%)
Red Delicious	34.58	34.79	0.21 (0.61%)	449 (93%)
Market Shares of the Honeycrisp ≥ 5 percent (96 Markets)				
Braeburn	67.08	67.48	0.39 (0.60%)	73 (76%)
Fuji	53.40	53.74	0.34 (0.64%)	81 (84%)
Gala	31.37	32.07	0.71 (2.23%)	95 (99%)
Golden Delicious	52.67	52.77	0.10 (0.19%)	70 (73%)
Granny Smith	41.92	42.15	0.24 (0.55%)	69 (72%)
Pink Lady/Cripps Pink	62.50	62.93	0.43 (0.69%)	74 (77%)
Red Delicious	33.44	34.00	0.56 (1.67%)	95 (99%)

Note: Price and C. Price represent the observed and the counterfactual price respectively. Both are averaged across retailer groups by sales quantity and deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period at 1982-84. The price change is the difference between these two prices and the percentage change in prices is presented in the associated parenthesis.

Table 3.9. Sales Quantity (Million Pounds) and Sales Revenue (Million Dollars)

			Sales Quantity					Sales Revenue				
Year	Num. of Markets	Num. of IRI Cities	Honeycrisp	Other Apples	Total	C. Total	Changes	Honeycrisp	Other Apples	Total	C. Total	Changes
2009	42	29	13.35	127.10	140.45	131.00	9.41	12.58	48.62	61.20	50.43	10.77
2010	61	38	21.06	194.80	215.86	201.20	14.63	17.77	78.09	95.86	81.03	14.84
2011	78	39	29.99	227.50	257.49	235.90	21.54	22.71	89.26	111.97	93.11	18.86
2012	82	38	29.67	267.10	296.77	277.10	19.73	28.17	112.19	140.36	117.53	22.83
2013	107	43	42.22	353.50	395.72	367.60	28.11	36.13	137.41	173.54	144.13	29.41
2014	111	43	47.66	347.50	395.16	362.10	33.06	42.46	128.37	170.83	135.42	35.41
Total			183.95	1517.50	1701.45	1574.90	126.48	159.82	593.94	753.76	621.65	132.12

Note: These results are based on the 481 markets where the market share of the Honeycrisp is greater than or equal to 1 percent. Other apples include all competing apple varieties. C. Total in sales quantity and sales revenue respectively represent the counterfactual quantity and revenue when the Honeycrisp is removed from the markets. The values of sales revenue are deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period at 1982-84.

Table 3.10. Decomposition of Welfare Changes for An Average Consumer (Cent per Pound)

Total Changes at Average in Consumer Welfare ($E[CV_i]$)	Changes from Observed Characteristics ($\delta_j + \mu_{ij}$)	Changes from Logit Error (ϵ_{ij})
Market Shares of the Honeycrisp ≥ 1 percent (481 Markets)		
3.14 (100.00%)	1.87 (59.55%)	1.27 (40.45%)
Market Shares of the Honeycrisp ≥ 5 percent (96 Markets)		
4.49 (100.00%)	3.18 (70.82%)	1.32 (29.18%)

Note: Average consumer welfare is estimated by the simulation form of $E[CV_i] = \int (u_i^{\text{with}} - u_i^{\text{without}})/\alpha_i \, dP(\epsilon)dP(D)dP(v)$ where $u_i = \max_j u_{ij}$ and ϵ is draw from the general extreme value distribution with shape parameter $\kappa = 0$, scale parameter $\sigma = 1$, and location parameter $\mu = 0$. The component ratios are presented in parentheses.

Table 3.11. Total Changes in Consumer Welfare (Million Dollars)

Year	Num. of Markets	Num. of IRI Cities	Change due to Increased Varieties	Change due to Decline in Prices of Competing Apples	Total Change in Consumer Welfare
2009	42	29	2.76 (91.09%)	0.27 (8.91%)	3.03 (100%)
2010	61	38	4.42 (92.28%)	0.38 (7.72%)	4.79 (100%)
2011	78	39	6.73 (92.45%)	0.54 (7.55%)	7.28 (100%)
2012	82	38	7.05 (91.56%)	0.66 (8.44%)	7.70 (100%)
2013	107	43	10.04 (91.11%)	0.98 (8.89%)	11.02 (100%)
2014	111	43	13.91 (91.51%)	1.29 (8.49%)	15.20 (100%)
Total			44.91 (91.60%)	4.12 (8.40%)	49.03 (100%)

Note: These results are based on the 481 markets where the market share of the Honeycrisp is greater than or equal to 1 percent. The values of consumer welfare are deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period at 1982-84. The component ratios are presented in parentheses.

Figure 3.1. Annual Sales of Honeycrisp (Million Pounds)

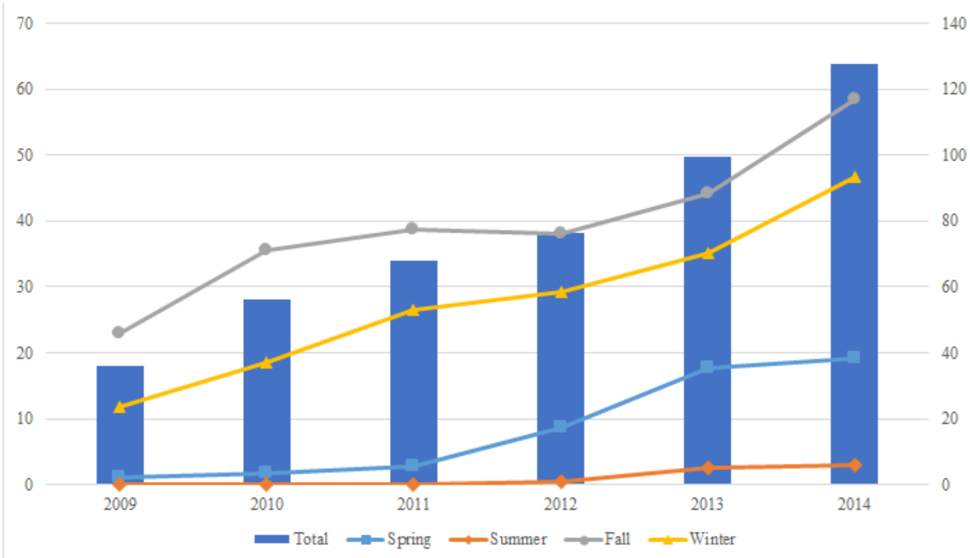


Figure 3.2. Map of the Cities in IRI data and Terminal Markets

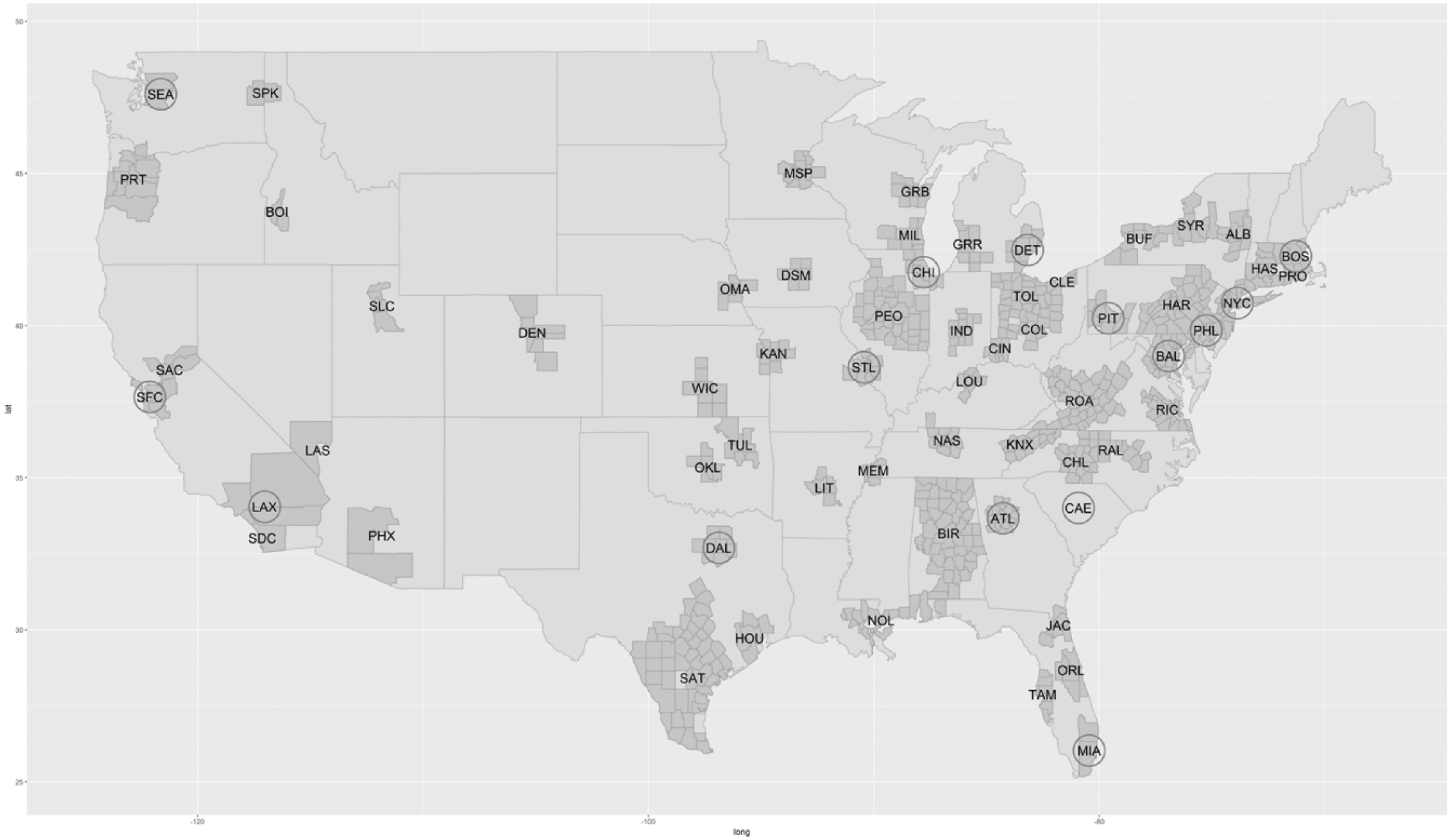
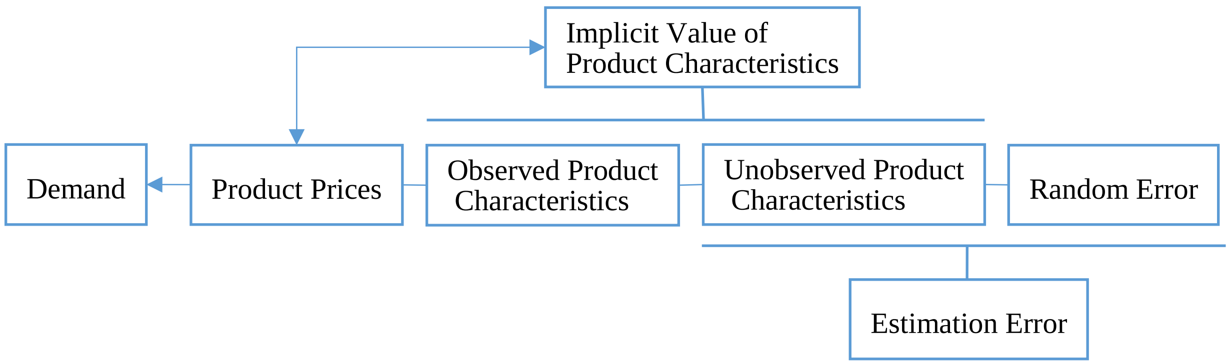


Figure 3.3. Correlation between Product Prices and Estimation Error



Note: Product prices represent the implicit value of all product characteristics, but not all product characteristics are included in the demand estimation. Therefore, the correlation between product prices and the estimation error, which contains the consumer valuation of unobserved characteristics, raises the problem of endogeneity.

Chapter 4. Food Retail Competition in Local Markets

4.1. Introduction

The food retail industry in the United States (U.S.) is dominated by a few retail chains. In 2012, there were 42,420 retail chains operating 66,343 stores, while the largest 20 retail chains earned a total of 315 billion dollars, accounting for 59 percent of the entire sales revenue in the food retail industry (U.S. Census Bureau 2015). Retailer concentration in local markets can be even higher. Typically, in a market with more than five million people, the majority of sales are captured by three to six retail chains, and the rest is shared by a fringe of small retailers (Ellickson 2013). Previous studies provide strong evidence on a positive correlation between retail price levels and the local market concentration (e.g., Cotterill 1986, 1999). Multi-store retail chains are able to internalize cross-effects between stores within the same chain, and hence enhance market power and result in high markups (Smith 2004).

The knowledge of food retail competition in local markets could inform public and business policy. In 2010, the U.S. Departments of Agriculture (USDA), Treasury, and Health and Human Services (HHS) launched the Healthy Food Financing Initiatives (HFFI) to increase food access in underserved communities through efforts such as developing and equipping grocery stores, small retailers, corner stores and farmers markets selling healthy food (U.S. Department of Health and Human Services 2017).[38] Food access is determined by both the accessibility and affordability of food. Ideally, opening a new store will directly increase the accessibility of food and raise competition in the local market and in turn reduce food prices in the local market. Nevertheless, stores in the local market might cooperate with their nearby rivals and thereafter exercise market power and maintain considerable profit margins. Kyureghian et al. (2012) find

[38] The HFFI received 125 million dollars from the USDA through the 2014 Farm Bill, 22 million dollars from the Treasury's Community Development Financial Institutions Fund in 2015, and a total amount of funding over 51.8 million dollars from the HHS between 2011 and 2016 (U.S. Department of Health and Human Services 2017).

most households fail to meet the USDA dietary guidelines because they cannot afford healthy foods. As a result, it is of importance to understand household valuations of store proximity and the extent to which retailer concentration drives up food prices in local markets.

The primary objective of this chapter is to examine store competition in local markets by addressing the heterogeneity in households' store choice sets, travel distances, and shopping baskets. A secondary objective is to evaluate the association between the profit margin and the number and composition of nearby rivals in local markets. To achieve these objectives, we develop a structural model that accounts for households' store choice and allows us to evaluate price elasticities and gross profit margins across stores and retail chains. The data used in this study are primarily from two sources—the Consumer Network household scanner data and the InfoScan retail scanner data—including rich and complementary information on household grocery shopping and weekly store sales. The joint use of these two data is novel to the research on the household choice of stores.

There is a sizable literature on food retail competition. Early studies using structure-conduct-performance models generally find a strong positive relationship between the market concentration and retail prices, and conclude that retailers with large market shares earn high profits by raising prices (e.g., Cotterill 1986, 1999). Bonanno and Lopez (2009) estimate a simultaneous equation system of prices and store service using store-level data on fluid milk purchases. They find that supermarkets differentiate themselves from their rivals through store service. A strand of the literature looks into retail competition across store formats. Hausman and Leibtag (2007) investigate the change in prices due to the entry of supercenters by estimating consumer choices between supercenters and traditional food outlets using household panel data. Their results suggest that consumers, especially from low-income households, benefits from lower food prices offered by the supercenters, and that the expansion of supercenters leads to a significant increase in consumer welfare. Volpe (2011) examines the performance (i.e., profitability and cost efficiency) of supermarkets and their competition with hypermarkets (e.g., Wal-Mart Supercenter). His results reveal that the competition with hypermarkets cannibalizes the profits of supermarkets. In fact, store competition exists not only across different formats and retail chains but also within the same retail chain. Cleeren et al. (2010) use an empirical entry model to examine the competition between discount stores and supermarkets in the German grocery industry. They find that the intra-format competition is more intensive than the inter-format competition. In another study, Smith (2004) estimates cross-elasticities between stores within the same retail chain. His results show that demerger for stores within the largest firms brings a price decrease of 2 to 3.8 percent in selected regions varying by market concentration, while a merger between the largest firms leads a price increase up to 7.4 percent. A market is defined in two dimensions—differentiated products and geographic areas. However, previous studies typically investigate the retail competition in the product space. A distinguishing feature of this study is that we look into this issue in both the product space and the geographic space by accounting for household-specific shopping baskets and travel distances, as well as their heterogeneous store choice sets.

This chapter is also related to the literature on determinants of households' store choice. The existing papers on food retail competition typically examine store substitution patterns by modeling the household choice of shopping stores. For example, Volle (2001) estimates a multinomial logit model to evaluate the impact of store promotions and loyalty on store choice using household behavioral data on grocery shopping in France and concludes that loyalty is a primary determinant. Including more variables of store characteristics, Solgaard and Hansen (2003) employ a random coefficient logit model to identify the determinants of household store choices. They find that price level, assortment, and distance determine the consumer choice, and that assortment is the most important factor driving the choice among different formats and the impact of distance varies with consumer demographics. In addition, their results show that there is no significant difference in quality and service across stores. Recent studies provide new empirical evidence by including both store and household characteristics in their analyses. Dong and Stewart (2012) develop a specialized multinomial logit model and find the consumer choice of stores for purchasing fluid milk heavily depends on store prices, promotion deals, as well as household demographics and purchase history. Kyureghian and Nayga Jr (2013) take a further step by integrating county-level retail availability information into Dong and Stewart's model to examine the impacts of local food retail environment on the household purchase of fruit and vegetables. They find that the availability of convenience stores drives up the household probability of purchasing fruit and vegetables.

Most abovementioned studies account for variables of household demographics, store average prices, and store characteristics in their models under an implicit assumption that different households visiting the same store will buy an identical basket. However, this implicit assumption is tenuous if household shopping baskets are substantially different from one another. We address the heterogeneity in households' store choice and shopping baskets by constructing prices for household-specific shopping baskets. Turolla (2016) is the only other economic study that accommodates this issue using the microdata on household grocery shopping in France. With limited information on the composition of household shopping baskets and their expenditures, Turolla develops a model to predict prices of household-specific shopping baskets that are, in turn, used in the estimation of the household store choice model. The author finds that retailers' market power in some local markets is simply due to a high level of concentration rather than the firms' anti-competitive practices, and concludes that the competition between large grocery stores is highly localized in France. A notable advantage of our study is that we are able to calculate prices of household-specific shopping baskets charged by different stores using both the household and retail scanner data.

This chapter is organized as follows. Section 2 gives us an overview of data and provides some preliminary evidence on the tradeoff between household travel distance and shopping basket price. Section 3 shows the analytical framework by introducing our structural models of demand and supply. Finally, we present and discuss the model results in Section 4, and conclude in Section 5.

4.2. Overview of Data

4.2.1. Data and Local Markets

Data in our analysis come from multiple sources. The primary data comprise the Consumer Network household scanner data and the InfoScan retail scanner data from January to December in 2016, both of which are collected by Information Resources, Inc. (IRI).[39] The household scanner data are derived from a large sample of nationally representative households who report their demographics, when they shopped, where they visited, what food products they purchased, and how much they spent. Household demographic data include household income level, size, and residence area, as well as the gender and the education level of household head, where a household head in this study is a household member who makes household decisions on shopping. The amount of money a household spent on each food product is measured by dollars at the Universal Product Code (UPC) level. For the sake of household privacy, two restrictions are imposed on the data. First, households only report names of retail chains they visited rather than the specific store locations. This limitation creates a challenge in modeling households' store choice that we discuss below. Second, a household's residence area can only be tracked by a census block, which is the smallest geographic unit defined by the U.S. Census Bureau. This restriction is less of a concern for this study since census blocks are informative enough to infer household residence in our analysis. In particular, we focus on the Minneapolis and St. Paul (MSP) Metropolitan Statistical Area (MSA), the 16th largest metropolitan area in the United States with more than 3.6 million

[39] The IRI data access is granted by the USDA Economic Research Service.

people. The household scanner data representatively sampled 3,329 households from 2,872 census blocks spreading out across 11 counties in the MSP MSA.[40]

The retail scanner data contain weekly sales information (i.e., sales revenue and quantity) of stores from the participating retail chains at the UPC level. Every store can be identified and linked to a specific address represented by the latitude and longitude coordinates. In the MSP MSA, there are 5 retail chains included in the data.[41] Retail chain 1, 2, and 3 report sales information at the store level, while retail chain 4 and 5 report at the Retail Market Area (RMA) level. An RMA is a geographic marketing area that is defined by the retail chain and might cross the boundaries of states and counties. In this study, the stores from retail chain 4 are all from the same self-defined RMA, so are the stores from retail chain 5. As a result, there is no variation in sales revenues and prices for stores from these two chains.

Figure 4.1 presents the location information of households and stores in our sample. Stores from retail chains are marked by different shapes of points, while households are labeled by plain dots. The information of household density at the city (i.e., subdivision) level is also presented in the figure by the degree of grey. The figure suggests that the majority of stores and households in the MSP MSA are located in the area whose boundary is denoted by the interior dash line. This area represents the core market in the MSA MSP, covering 7 counties—Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington—in Minnesota. Aware of households who visit the stores on the boundary from outside but close to the core market, we extend the area of the core market outward by 5 kilometers (km). To that end, our sample includes all households and stores located in the area, the boundary of which is denoted by the outer dash line.

[40] These 11 counties consist of 10 counties (i.e., Anoka, Carver, Chisago, Dakota, Hennepin, Ramsey, Scott, Washington, and Wright) in Minnesota and 1 country (i.e., St. Croix) in Wisconsin.

[41] Because of the confidentiality agreement, we are not allowed to share the names of retail chains in this study.

There are 1280 households in the sample. Table 4.1 displays the summary statistics of household demographics. It shows 21 percent of sampled households are headed by a male, 59 percent are married, and 87 percent are white. More than half of sampled household heads hold a bachelor's or above degree. Approximately 75 percent of the households are staying in their own houses. Half of the households are employed, and 16 percent of them are partially employed. About 29 percent of households have at least one kid. Most sampled households consist of 2 individuals with an average size of 3. Compared to the median household income in the United States (55,322 dollars in 2016), households residing in the MSP MSA are wealthier. Specifically, only 10 percent of sample households earn an income less than 25,000 dollars a year and 46 percent of sample households have an annual income of more than 70,000 dollars.

The household scanner data include information on household shopping trips and products they purchased at the UPC level. A household might shop in multiple stores during one shopping trip and shop multiple times in the same store in a given month. For the subsequent analysis, we aggregate household shopping trip data at the monthly level by store and food category and focus on shopping trips made to households' primary stores. A primary store for a household in a month is the store at which the household spends the largest portion of its monthly grocery expenditure. A household on average spends 67 percent of its monthly expenditure in its primary store. Overall, 62 percent of household primary stores are from the five retail chains included in the retail scanner data.

Table 4.2 shows the average household expenditure shares of 55 food categories in two groups: perishable product items and processed product items. The most consumed food category of perishable product items is Fruit, accounting for 5.42 percent of an average monthly household grocery expenditure, followed by Vegetables 3.59 percent, Beef 2.81 percent, Deli Prepared 1.57 percent, and Pork 1.41 percent. As for the processed product items, the most consumed food category is Dairy with 12.86 percent, followed by Snacks 6.35 percent, Frozen Meals 5.19 percent, Refrigerated Meats 5.16 percent, and Meals 5.03 percent. On average, 20 percent of monthly household expenditures are spent on perishable product items. These household expenditure patterns largely mirror the average shares of store revenues by food category, which are derived from the retail scanner data and presented in Table C1.

Table C2 shows the average store prices by food category. In particular, the average store price p_{jct} of food category c in store j at period t is calculated by

$$p_{jct} = \sum_{k=1}^{K} \omega_{jckt} p_{jckt} \tag{24}$$

where ω_{jckt} and p_{jckt} are the revenue share and the unit price of product item k in food category c in store j at period t.

A distinct contribution of this study is that we examine retail competition in local markets, which are identified by households' willingness to travel for shopping (WTTFS). A household's WTTFS is a scalar, implying the size of a local market for a store. A local market is an area of a circle centered on a store with a radius equal to the assumed household WTTFS. Table 4.3 presents the summary statistics of the data on household travel distance to stores, the unit price of the household basket, and the size of household store choice sets under different assumptions of household WTTFS, ranging from 10 km to 20 km. Note that the size of a household store choice set is positively related to the household WTTFS. Panel A shows the distribution of household travel distances to stores in household choice sets. The household travel distance to a store is the Euclidian distance between the household home address and the store address.[42] Suppose households are equally likely to visit every store in their choice sets. The expected household travel distance to a store on average increases from 6.49 km to 12.63 km, as the household WTTFS raises from 10 km to 20 km. The 2009 National Household Travel Survey, conducted by the U.S.

[42] For privacy protection, households are only able to be tracked at the census block level. Therefore, a household home address is regarded as the centroid of the census block in which the household resides.

Department of Transportation Federal Highway Administration, reports the annual average distance of a personal trip for shopping was 6.5 miles, equivalent to 10.4 km (Santos et al. 2011). This is close to 10.23 km, the average household travel distance to a shopping store when households are willing to travel as far as 16 km for shopping. Therefore, we use 16 km as the benchmark distance for household WTTFS. Robustness checks are conducted throughout the subsequent analyses. As the household WTTFS doubles from 10 km to 20 km, the sample size (i.e., the overall number of stores in households' choice sets) increases by more than three times, from 112,932 to 372,067 observations. Panel B of Table 4.3 presents the summary statistics of the unit price of the household basket. The unit price p_{hjt} of household h's shopping basket in store j at period t is given by

$$p_{hjt} = \sum_{c=1}^{C} \omega_{hct} p_{jct} \tag{25}$$

where ω_{hct} is the household h's expenditure share of food category c at period t and p_{jct} is the average price of food category c in store j at period t. The statistics suggest the distribution of the unit price of household shopping basket does not vary much with the household WTTFS. According to Equation (25), the unit price of a household shopping basket is the unit price of a composite good by construction. Panel C of Table 4.3 shows the change in the size of a household choice set. As the household WTTFS increases from 10 km to 20 km, the median number of stores that are included in a household choice set inflates from 11 to 39.

Suppose households are willing to travel as far as 16 km. Then, the local market of a store is an area of a circle centered on the store with the radius of 16 km. The union of all households' choice sets contains 119 stores in the study area. The union of local markets overlaps 141 cities from 10 counties in Minnesota (MN) and 1 county in Wisconsin (WI). Note that each city is a subdivision of a county and it is a small administrative area in the United States. Table 4.4 displays the summary statistics for these cities. As shown in Figure 4.1, Hennepin and Ramsey County are the top two most populated counties in the study area. On average, a city in Hennepin and Ramsey County has 12,258 and 10,912 households with the density of 400 and 461 households per squared kilometer, respectively.

Table 4.5 presents the summary statistics of total revenue, store price, and market shares by retail chain and by store. Retail chain 1 is the leading retailer operating 44 stores in the study area with the total revenue of 91.7 million dollars in 2016, followed by retail chain 2 operating 22 stores with 56.3 million dollars, and retail chain 5 operating 20 stores with 41.1 million dollars. A store price is an average price across food categories weighted by its revenue; that is,

$$p_{jt} = \sum_{c=1}^{C} \omega_{jct} p_{jct}. \tag{26}$$

where ω_{jct} and p_{jct} are the sales revenue and the price of food category c in store j at period t. On average, stores from retail chain 5 have the lowest store prices. The market share of a retail chain is the quantity sold in the retail chain divided by the total number of potential quantity in the study area.[43] The total number of quantity sold in a retail chain is defined as the total revenue of the retail chain divided by the average store price of the retail chain. Following the previous literature, the total number of potential quantity is a product of the total number of households and a constant, which in this study is the ratio of the monthly household average grocery spending over the average store price across retail chains. The use of potential quantity allows us to model a scenario in which the current market could be enlarged. According to the USDA moderate food plan, the monthly average cost of food at home is 605 dollars for a household of two adults, 874 dollars for a household of a couple with two children between 2 and 5 years old, 1043 dollars for a household of a couple with two children between 6 and 11 years old in December 2016 (USDA 2017b).[44]

[43] The number of quantity in our context refers to the number of shopping baskets.

[44] The USDA Center for Nutrition Policy and Promotion releases a monthly online report for the cost of food on the page: https://www.cnpp.usda.gov/reports-publications.

Since the average size of households in the sample is about 3, we assume the monthly household average grocery spending is 650 dollars.[45]

Table 4.5 shows the overall market share of the leading retailer, retail chain 1, is 13.02 percent in the study area, which is 5.54 and 7.70 percent greater than the market share of the second and third largest retailer, respectively. Analog to the calculation of the market share of a retail chain, we also compute the market share of a store within the chain. Table 4.5 shows that the average market share of a store from retail chain 1 is 1.10 percent, which is much greater than the number, the overall market share of retail chain 1 over the total number of stores within retail chain 1. This suggests the market power of an individual store might be negligible in a large market but significant in a local market. It lends support to the argument of local markets.

4.2.2. Preliminary Evidence on Tradeoff between Travel Distance and Shopping Basket Price

The household scanner data only include household choices of retail chains instead of specific store locations. Before turning to the model, we examine the extent to which we can infer household choices of retail chains from observed data. Table 4.6 presents the summary statistics of household travel distance and shopping basket price based on alternative subsamples. In Scenario 1, the data is first restricted to a sample that contains the choice of the closest store(s) for each household and then restricted to a subsample that contains the choice of the closest store with the lowest price. Accordingly, the average household travel distance is 2.53 km, the average unit price of household shopping basket (i.e., the average unit price of a composite good) is 4.01 dollars, and in the selected subsample there is 39 percent of the choice options containing the stores from the retail chains that

[45] The results from the subsequent analysis are robust as we change the grocery spending from 625 dollars to 725 dollars by every 25 dollars.

households visit. In turn, the data in Scenario 2 is first restricted to a sample that includes the choice of the store(s) with the lowest price for each household and then restricted to a subsample that contains the choice of the closest store with the lowest price. Accordingly, the average household travel distance increases to 9.66 km, while the average unit price of household shopping basket goes down to 3.67 dollars and the percentage of the selected choices with the stores from the retail chains that household visit declines to 25 percent. These statistics imply that households could find a further store with a lower average price for their shopping but they are willing to visit a closer store and pay a higher price instead. The same implications are found in Scenario 3 and 4, where we only focus on the data including the household choices with the stores from the retail chains that they visit. In sum, the tradeoff between household travel distance and shopping basket price motivates our research on food retail competition in local markets.

4.3. Analytical Framework

4.3.1. The Demand Model

Discrete choice models are widely used in the estimation of demand in the empirical IO studies and are capable of investigating the consumer choices of differentiated products where each product has a constant price over all households. In our context, shopping stores are the differentiated products in the household choice sets. Analog to the existing literature, the variable of store average price can be included to model the household choice of stores, implicitly assuming that households visiting the same store buy an identical shopping basket and hence pay the same price. In fact, households may decide to visit a store depending on the composition of their shopping baskets. Using the household and retail scanner data, we are able to relax this assumption by measuring the price charged by every store in a household's choice set for its specific basket. Furthermore, another variable of household travel distance to a store is also included to capture the tradeoff between household travel distance and shopping basket price.

Let households be indexed by $h = 1, \ldots, H$ and stores in the household h's choice set J_h by $j = 1, \ldots, |J_h|$ where $|J_h|$ is the size of household h's choice set. The indirect utility that a household h achieves from shopping in store j at period t is given by

$$V_{hjt} = \alpha_h p_{hjt} + \lambda_h dist(L_h, L_j) + \boldsymbol{X}_h' \boldsymbol{\beta} + \boldsymbol{X}_j' \boldsymbol{\gamma} + \boldsymbol{D}_t' \boldsymbol{\phi}, \tag{27}$$

where p_{hjt} is the household-store price at period t defined in Equation (25), $dist(L_h, L_j)$ is the household h's travel distance to store j, $\boldsymbol{X}_h$ is a vector of household demographics, $\boldsymbol{X}_j$ is a vector of store characteristics, $\boldsymbol{D}_t$ is a vector of month dummies, α_h and λ_h are scalar parameters to be estimated, and $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\phi}$ are the conformable vector of parameters to be estimated. The utility from choosing an outside option ($j = 0$) is normalized to be zero (i.e., $V_{h0t} = 0 \ \forall h, t$). To account for household heterogeneous tastes of price and travel distance, we allow the parameters α_h and λ_h to vary with household demographics $\boldsymbol{Z}_h$ and unobserved factors $\boldsymbol{v}_h$ such that

$$\begin{pmatrix} \alpha_h \\ \lambda_h \end{pmatrix} = \begin{pmatrix} \alpha \\ \lambda \end{pmatrix} + \boldsymbol{\Pi} \boldsymbol{Z}_h + \boldsymbol{\Sigma} \boldsymbol{v}_h, \tag{28}$$

where the first component $(\alpha, \lambda)'$ represents the average taste of price and travel distance over households, the second component $\boldsymbol{\Pi} \boldsymbol{Z}_h + \boldsymbol{\Sigma} \boldsymbol{v}_h$ represents household-specific deviations from the average, and $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$ the conformable matrices of parameters to be estimated. To complete the model, an independent and identically distributed error term ϵ_{hjt} is included to capture the unobserved disturbance in the valuation of household h's choice of store j for grocery shopping. The model hence can be written as

$$U_{hjt} = V_{hjt}(\boldsymbol{\theta}_h) + \epsilon_{hjt}, \tag{29}$$

where $\boldsymbol{\theta}_h = [\alpha_h, \lambda_h, vec(\boldsymbol{\beta}), vec(\boldsymbol{\gamma}), vec(\boldsymbol{\phi})]$ is a vector of all parameters. Following the literature on discrete choice models, ϵ_j is assumed to be drawn from the Type I Extreme Value distribution. The probability of household h choosing store j at period t is conditional on $\boldsymbol{\theta}_h$, given by

$$s_{hjt}(\boldsymbol{\theta}_h) = \frac{\exp V_{hjt}\,(\boldsymbol{\theta}_h)}{\sum_{k\in J_h} V_{hkt}\,(\boldsymbol{\theta}_h)}. \tag{30}$$

This corresponds to the choice probability of a mixed logit model. It will reduce to the choice probability of a multinomial logit model if the parameter matrices $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$ are restricted to be zero.

Recall that households only report the names of retail chains rather than the specific locations of shopping stores. To address this restriction, we have to match the information on the choice probabilities at the retail level. Let retail chains be indexed by $r = 1, \ldots, R$. The probability of household h visiting retail chain r at period t is

$$s_{hrt}(\boldsymbol{\theta}_h) = \sum_{j\in J_h(m)} I_r(j) s_{hjt}(\boldsymbol{\theta}_h), \tag{31}$$

where $J_h(m)$ is the household h's choice set of stores within m kilometers of its home address, and $I_r(j)$ is an indicator function such that $I_r(j) = 1$ if store j is operated by retail chain r and 0 otherwise. Let $r(h,t)$ denote the retail chain that household h reports at period t. Conditional on $\boldsymbol{\theta}_h$, the probability of household h's observed sequence of choices in the data is given by

$$s_h(\boldsymbol{\theta}_h) = \prod_t s_{hr(h,t)t}(\boldsymbol{\theta}_h). \tag{32}$$

The unconditional probability of household h's sequence of choices is

$$P_h(\boldsymbol{\theta}) = \int s_h(\boldsymbol{\theta}_h) f(\boldsymbol{\theta}_h \mid \boldsymbol{\theta}) d\boldsymbol{\theta}_h, \tag{33}$$

where $\boldsymbol{\theta}$ is a vector of parameters in the probability density function for $\boldsymbol{\theta}_h$. Therefore, to evaluate the probability of household h's observed sequence of choices, we need to draw a vector $\boldsymbol{\theta}_h$ from the distribution with a given $\boldsymbol{\theta}$. For repeated draws indexed by $n = 1, \ldots, N$, the average of $s_h(\boldsymbol{\theta}_h)$ over $\boldsymbol{\theta}_h$ is an approximation to $P_h(\boldsymbol{\theta})$, that is,

$$SP_h(\boldsymbol{\theta}) = \frac{1}{N}\sum_{n=1}^{N} s_h(\boldsymbol{\theta}_h^n), \tag{34}$$

where $\boldsymbol{\theta}_h^n$ is the value of $\boldsymbol{\theta}_h$ in the n-th draw from $f(\boldsymbol{\theta}_h \mid \boldsymbol{\theta})$ and $SP_h(\boldsymbol{\theta})$ is the simulated unconditional probability of household h's sequence of choices. Therefore, the simulated log-likelihood function is given by

$$SLL(\boldsymbol{\theta}) = \sum_{h=1}^{H} \log\big(SP_h(\boldsymbol{\theta})\big). \tag{35}$$

Due to the highly nonlinear structure of the simulated log-likelihood function in Equation (35), we use the corresponding gradient function in the estimation to reduce the computational burden. We employ Halton draws for simulation instead of independent random draws to increase accuracy. A total of 100 Halton draws are generated for the model estimation (e.g., Petrin and Train 2010; Turolla 2016).[46]

The substitution patterns across stores are characterized by the stores' price elasticities, which measure the changes in stores' market shares (sales quantities) in response to the changes in store prices. For example, store j's elasticity with respect to store k's price is given by,

$$\frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \dfrac{p_{kt}}{s_{jt}} \displaystyle\int \alpha_h s_{hjt}\big(1 - s_{hjt}\big) f(\boldsymbol{\theta}_h \mid \boldsymbol{\theta}) d\boldsymbol{\theta}_h & \text{if } j = k \\ -\dfrac{p_{kt}}{s_{jt}} \displaystyle\int \alpha_h s_{hjt} s_{hkt} f(\boldsymbol{\theta}_h \mid \boldsymbol{\theta}) d\boldsymbol{\theta}_h & \text{otherwise} \end{cases}. \tag{36}$$

This equation implies that store j's own-price elasticity depends on its own market share but its cross-price elasticities vary with rivals' market shares. Compared to a standard multinomial logit model and a nested logit model, a mixed logit model relaxes the assumption of the independence of irrelevant alternatives to a large extent and provides more accurate estimates of store substitution patterns.

[46] The performance of 100 Halton draws is better than that of 1000 independent random draws (Bhat 2001).

4.3.2. The Supply Model

To examine the competition of stores in local markets and investigate the variation of gross profit margins across stores, we need to know their marginal costs, which are unobserved from the data. In the empirical IO studies, the marginal costs are typically inferred by the firm's pricing conditions derived from an oligopolistic competition model. In the model, a firm is assumed to maximize its profit by setting prices in each period. The first order conditions are used to characterize the equilibrium solution to this profit maximization problem. Then, marginal costs can be recovered using the first order conditions together with the demand estimates. Consequently, profit margins can be attained.

We estimate stores' gross profit margins in three alternative scenarios of retail competition in local markets. The first scenario is the most competitive case, in which stores are assumed to compete against each other and prices are set by individual stores. The second scenario is motivated by the observation from Figure 4.1 in which some stores are located on the same site. In this scenario, another store in a place within 1 km from a store is considered as its rivals on the same site and all stores are assumed to collude with their rivals in the same site only. In the last scenario, stores within the same retail chain are assumed to coordinate together and maximize their joint profits. For the purpose of illustration, we present the supply model in the last scenario.

Suppose a retail chain r is operating a collection J_r of stores in our sample. The expected profit function of retail chain r at period t is given by

$$\Pi_{rt} = \sum_{j \in J_r} \left(p_{jt} - c_{jt}\right) M_{jt} s_{jt}(\boldsymbol{p}_{rt}, \boldsymbol{p}_{-rt}) - C_{jt}, \tag{37}$$

where p_{jt} is the store price defined by Equation (26), c_{jt} is the associated marginal cost, M_{jt} and s_{jt} are the local market size and market share for store j at period t, respectively. The price vector $\boldsymbol{p}_{rt}$ and $\boldsymbol{p}_{-rt}$ represent a collection of store prices for stores from retail chain r and others, respectively. The fixed cost C_{jt} for store j at period t is a constant. As a result, stores' first order conditions to this profit maximization problem for retail chain r is

$$s_j(\boldsymbol{p}_{rt}, \boldsymbol{p}_{-rt}) + \sum_{k \in J_r} (p_k - c_k) \frac{\partial s_k(\boldsymbol{p}_{rt}, \boldsymbol{p}_{-rt})}{\partial p_j} \frac{M_k}{M_j} = 0, \forall j \in J_r. \tag{38}$$

In contrast to the previous studies where every store competes against rivals in the entire market, Equation (38) contains a term M_k/M_j reflecting the size ratio of different local markets and this term will reduce if the sizes oftwo local markets are the same (i.e., $M_k = M_j$ for $k \neq j$). Denote Δ as the response matrix of store market share with respect to prices, where $\Delta_{kj} = \left(\partial s_k/\partial p_j\right) \cdot \left(M_k/M_j\right)$ for all k and j. Then, the first order condition in vector notation is

$$\boldsymbol{s}(\boldsymbol{p}) + [\boldsymbol{T} \otimes \Delta(\boldsymbol{p})](\boldsymbol{p} - \boldsymbol{c}) = 0, \tag{39}$$

where $\boldsymbol{p}$ is a vector of observed store prices, $\boldsymbol{c}$ is a vector of marginal costs to be estimated, $\otimes$ represents the Kronecker product, and $\boldsymbol{T}$ is an ownership matrix with $\boldsymbol{T}_{jk} = 1$ if $j, k \in J_r$ and $\boldsymbol{T}_{jk} = 0$ otherwise. The differences between store prices and associated marginal costs are hence given by

$$(\boldsymbol{p} - \boldsymbol{c}) = -[\boldsymbol{T} \otimes \Delta(\boldsymbol{p})]^{-1} \cdot \boldsymbol{s}(\boldsymbol{p}). \tag{40}$$

Therefore, the stores' gross profit margins can be calculated by $(\boldsymbol{p} - \boldsymbol{c})/\boldsymbol{p}$.

4.4. Results

4.4.1. The Demand Model Estimates

The household store choice model is essentially a revised mixed logit model of demand for stores that are characterized by differentiated shopping baskets. We estimate this model using the method of simulated maximum likelihood. Table 4.7 shows the estimated parameters in the demand model under different assumptions of household WTTFS. Overall, the magnitude and significance of estimates are robust for the mean (homogeneous) part of random coefficients rather than the deviation (heterogeneous) part. This is probably because our model is estimated using aggregate information on households' choice of retail chains and, as household's WTTFS increases, the variation in household choice sets declines and subsequently reduces the model capability in identifying parameters in the deviation part of random coefficients. The last column presents a model specification that is often used in previous studies where household valuations of store proximity are not taken into account. For ease of interpretation, we focus on the benchmark model in which households are willing to travel as far as 16 km for shopping. Section 2.1 includes the detail discussion on household WTTFS in this study.

The estimated coefficients for household shopping basket price and travel distance are negative and statistically significant, suggesting that households bear a disutility from higher prices and longer travel distances to stores. In particular, the price estimates show an inverse U-shape relation between the magnitude of the price coefficient and the level of household income. The price coefficient is allowed to differ by household income. Compared to households with income less than 25,000 dollars, households with income more than 50,000 dollars are less sensitive to price. The average price coefficient for households with income between 50,000 and 70,000 dollars is the mean part of –0.394, homogenous over households, plus the differential of 0.142, for an estimate of –0.252. The estimates of the average price coefficients for households with income between 70,000 and 100,000 dollars and with income more than 100,000 dollars are –0.197 and –0.228, respectively. The average distance coefficient is an estimate of –0.369, which is less than the estimates of the average price coefficients for the majority of households. This implies that travel distance for shopping generates larger disutility than shopping basket price on the household valuation of store choice, which is in line with the preliminary evidence. Figure 4.2 presents the distributions of household willingness to pay (WTP) for travel distance by income level, measured by λ_h/α_h. In general, it shows a household's WTP for travel distance is increasing in its income level. However, for households with income above 100,000 dollars, the distribution of their WTP for travel distance is not in the rightmost place. Instead, it moves toward left and the average WTP for travel distance for households with income above 100,000 dollars is lower than the average level for households with income between 70,000 and 100,000 dollars. It is possible in the study area that households in the highest income group would spend more on food away from home and, thus, may not care about grocery shopping as much as households in other income levels.

Moreover, households headed by a female are more likely to visit the sampled stores than households headed by a male, while married household heads are less likely to shop in the sampled stores than single household heads. Compared to household heads with less than high school degrees, household heads with high school degrees are more willing to shop in the sampled stores whereas household heads with graduate (or professional) degrees are less. It is worth noting that the sampled stores in this study are either grocery or mass merchandise stores. Probably, in the study area, households headed by married people and those with graduate degrees prefer a club store, such as Costco, which is not sampled in our data. The employment status of a household head does not play a significant role in grocery shopping, while the size of a household does. Compared to single households, households consisting of 2 to 4 people are more willing to shop in the sampled stores.

In addition, the fixed effects of time and retail chains are captured by associated binary variables. Note that except for store prices, there are no more variables of store characteristics included in the model. Using store sales information, we are able to generate a continuous variable measuring the assortment strategy of a store (i.e., the total number of products sold in the store) and a binary variable indicating if a store has a salad bar or a deli department. However, these two variables and the retail chain dummies are highly collinear since stores from the same retail chain are usually designed in the same operational format. Therefore, the data do not enable us to capture the variation in store characteristics other than prices.

4.4.2. Price Elasticities and Profit Margins

Using the demand estimates, we can calculate stores' price elasticities and their gross profit margins. Table 4.8 presents the summary statistics of stores' price elasticities by retail chain and by county. Overall, the mean of own-price elasticities is elastic, while the mean of cross-price elasticities is relatively small and inelastic. This finding is similar to the result in Turolla (2016), suggesting a short-run change in prices might not affect the store market shares. Moreover, we find that, on average, stores in the counties with a larger population and more rivals have higher own- and cross-price elasticities than their counterparts in the counties with a smaller population and fewer rivals. For example, the average own-price elasticities over stores in Hennepin and Ramsey achieve –3.843 and –4.953, respectively, which are the largest two values in magnitude.

Next, we calculate stores' gross profit margins for the selected 106 stores, all of which have elastic own-price elasticities over the study period. Table 4.9 displays the summary statistics of gross profit margins computed under three alternative pricing strategies. The first column shows the distribution of these selected stores by retail chain and by county, and the number of rivals within the same site. The table shows that 40 percent of the selected stores are located on a site where there exists at least one rival. The lower bound of the gross profit margins vary by retail chain and by county, ranging from 29 percent up to 48 percent. Under the second pricing strategy, stores from retail chain 3 and stores in Washington County, on average, have the largest increases in their gross profit margins. Under the third pricing strategy, larger retail chains attain higher gross profit margins. In particular, stores from retail chain 1 on average attain the largest gross profit margin of 59.89 percent, followed by stores from retail chain 2, 5, 3, and 4. This ranking is the same as the ranking of retail chains in terms of market shares. Overall, regardless of pricing strategies, stores in Hennepin and Ramsey County—the counties with the top two highest store density (i.e., the number of stores per km^2)—have the lowest gross profit margins. This implies that the increase in local competition among stores will inevitably reduce their gross profit margins.

Last but not least, we are interested in the robustness of the estimates of stores' price elasticities and gross profit margins with respect to the assumption of household WTTFS. Particularly, we check the changes in these estimates in six different scenarios. In the first five scenarios, we assume households are willing to travel a certain distance for shopping, ranging from 10 to 20 km. In the last scenario, we follow the previous studies and assume households are willing to shop in every store in the study area regardless of travel distance. Figure 4.3 displays the distributions of the estimated stores' price elasticities in different scenarios. The boxplots in Figure 4.3 reflect that the medians of estimated own-price elasticities for stores from retail chain 1, 4, and 5 decrease in general as household WTTFS increases. The bottom right boxplot shows the overall distribution of the estimated elasticities for all stores, suggesting that compared to the first five scenarios, the median of estimated stores' own-price elasticities in the last scenario will be overestimated by 17.7 to 31.4 percent. Similarly, Figure 4.4 includes the boxplots of stores' gross profit margins. The bottom right boxplot specifically suggests that regardless of the assumed distance that households are willing to travel for shopping, the median of the estimated gross profit margins for all stores are underestimated by 9.5 to 17.4 percent if we fail to account for shopping in local markets.

4.4.3. Regression Analysis of Store Competition in Local Markets

Accordingly, we investigate how profit margins will vary with the competitive environment in local markets. In particular, we regress stores' gross profit margins on the number of nearby rivals and the demographics of the local population, that is, the local population percentages of marriage and different education and household income levels. The nearby rivals of a store are defined as the rivals within a certain distance away from the store.

For purpose of illustration, we present the regression results of stores' gross profit margins on the number of nearby rivals within 5 km in Table 4.10.[47] With the control variables of population demographics in local markets, we compare the estimates of interest within and across models in different specifications. Column (i) to (iii) display the regression results of gross profit margins in the first scenario that stores compete against each other, column (iv) to (vi) in the second scenario that stores collude within the same site, column (vii) to (ix) in the third scenario that stores collude within the same retail chain. On average, an increase of one rival within 5 km of a store is associated with a 2.4, 1.8, and 1.6 percent decrease in its gross profit margin in three alternative scenarios, respectively. To differ the rival effects from different retail chains, we replace the variable of the number of nearby stores with the number of nearby stores from the same retail chain and other retail chain, as well as a collection of the numbers of nearby stores from retail chain 1 to 5. Overall, most of these coefficients are negative and statistically significant, suggesting the gross profit margin of an average store will be cannibalized by its nearby rivals.

Moreover, we find that in the first two scenarios, the decrease in the gross profit margin for an average store is mainly from the competition between nearby rivals from the same retail chain in the local market. However, in the third scenario, the decrease is primarily from the competition between nearby rivals from other retail chains. The results reveal that an additional rival from other retail chain within 5 km of a store is related to a decrease of 1.8 percent in the store's gross profit margin. In this scenario where stores collude within the retail chain, there is a positive association between a store's gross profit margin and the presence of another store from the same retail chain. This association becomes statistically significant when the definition of nearby store distance is above 5 km. This implies that, in the third scenario, the distance between

[47] For robustness check, the regression results with different definitions of nearby stores are presented in the Appendix C.

each store from the same chain ought to be above, at least, 5 km to increase the joint profits of the retail chain.

In addition, the results show that, as the definition of the distance of nearby store increases, the association between a store's gross profit margin and the number of its nearby rivals becomes smaller but still significant. The results also show that the signs of the estimated coefficients for the number of nearby stores remain negative. For example, in Table C4, when nearby rivals are defined as other stores within 10 km, an increase in the number of nearby rivals will be related to a decrease of 1.1, 1.1, and 0.7 percent in the gross profit margin in three alternative scenarios, respectively.

4.5. Conclusion

We investigate food retail competition in local markets by modeling households' store choice in a structural framework that allows us to account for the heterogeneity in household choice sets of stores, travel distances, and shopping baskets. The IRI household and retail scanner data are jointly used in this study. Accordingly, a revised mixed logit model has been developed to accommodate the restriction in the household scanner data where households only report the names of the retail chains they visited rather than the locations of shopping stores. Using the estimated model, we are able to calculate stores' price elasticities and hence recover their gross profit margins under alternative retail pricing strategies that differ by the degree of coordination across stores. The results show stores' own-price elasticities and gross profit margins will be over- and under-estimated, respectively, if we neglect household willingness to travel for shopping. This finding strongly suggests the importance of our research on retail competition in local markets.

In particular, we focus on the Minneapolis and St. Paul Metropolitan Statistical Area with the household and retail data aggregated by month from January to December in 2016. We find the price of the household shopping basket and the household travel distance are negatively related to the household valuation of grocery shopping. Given a fixed level of household utility, there is an evident tradeoff between the price of household shopping basket charged by a store and the travel distance to it. The estimates of the household store choice model suggest that households prefer to visiting closer stores at expense of paying their shopping baskets at higher prices. The price coefficient implies that, as the level of household income increases, they will be less responsive to the change in the prices of their shopping baskets. The mean parts of random coefficients for store price and travel distance are robust when the maximal distance that households are willing to travel for shopping varies from 10 to 20 km.

Moreover, the average price elasticity over stores in an area reflects the local competitive environment. We find that, on average, stores in the counties with a larger population and more rivals have higher price elasticities than those in the counties with a smaller population and fewer rivals. Specifically, the results show that the average own-price elasticities over stores in Hennepin and Ramsey, the core counties made of the MSP MSA, achieve the top two largest values in the magnitude of –3.848 and –4.939, respectively.

Finally, regardless of alternative pricing strategies, the gross profit margin of a store is negatively associated with the intensity of local competition, characterized by the number of nearby rivals in the local market. Particularly, with the control of population demographics, a store in a local market with one more rival within its 5 km, on average, suffers a decrease of 1.6 to 2.4 percent in its gross profit margin depending on different pricing strategies. If stores compete against each other or collude with rivals in the same site, the decrease in the gross profit margin of a store is primarily attributed to the competition of nearby rivals from the same retail chain. If stores coordinate within the same retail chain, the gross profit margin of a store could, however, increase when another store from the same retail chain is opened in a place, at least, more than 5 km from the store. In other words, the joint profit of a retail chain is able to increase if the minimal distance between any two stores is, at least, more than 5 km.

Table 4.1. Summary Statistics of Household Demographics

Variable	Explanation	Mean	SD	Min	Max
Male	Household head is male	0.21	0.41	0	1
Married	Household head is married	0.59	0.49	0	1
White	Household head is white	0.87	0.34	0	1
High School	Household head has high school degree	0.17	0.38	0	1
Some College	Household head has some college or associate’s degree	0.30	0.46	0	1
College	Household head has bachelor’s degree (Reference)	0.39	0.49	0	1
Graduate	Household head has graduate or professional degree	0.14	0.35	0	1
Kid	Household has at least 1 kid aged less than 18	0.29	0.45	0	1
Owned House	Household owns the house	0.75	0.43	0	1
Fully Employed	Household head is fully employed	0.50	0.50	0	1
Partially Employed	Household head is partially employed	0.16	0.37	0	1
Household Size = 1	Household is made up of 1 individual (Reference)	0.27	0.44	0	1
Household Size = 2	Household is made up of 2 individuals	0.38	0.49	0	1
Household Size = 3	Household is made up of 3 individuals	0.14	0.35	0	1
Household Size = 4	Household is made up of 4 individuals	0.14	0.34	0	1
Household Size ≥ 5	Household is made up of more than 5 individuals	0.07	0.26	0	1
Household Income < $25k	Household income is less than $25,000 (Reference)	0.10	0.31	0	1
Household Income ≥ $25k and < $50k	Household income is greater than or equal to $25,000 but less than $50,000	0.26	0.44	0	1
Household Income ≥ $50k and < $70k	Household income is greater than or equal to $50,000 but less than $70,000	0.18	0.38	0	1
Household Income ≥ $70k and < $100k	Household income is greater than or equal to $70,000 but less than $100,000	0.23	0.42	0	1
Household Income ≥ $100k	Household income is greater than or equal to $100,000	0.23	0.42	0	1
Sample Size	Number of Sampled Households	1280			

Note: There are 1280 households sampled in the study area. Household demographic variables in the household scanner data are categorical, none of which is continuous.

Table 4.2. Average Household Expenditure Share by Food Category

Perishable Product Items		Processed Product Items			
Food Category	Share (%)	Food Category	Share (%)	Food Category	Share (%)
Baked Goods	0.95	Baby Food	0.45	Juices	1.31
Beef	2.81	Bakery	3.58	Liquor	2.75
Beef/Veal	0.05	Baking	3.38	Meals	5.03
Chicken	1.09	Breakfast	2.74	Non-Fruit Drinks	0.12
Deli Cheese	0.32	Candy	2.74	Other Frozen	0.00
Deli Meat	1.01	Carbonated Soft Drinks	2.81	Other Refrigerated	0.04
Deli Prepared	1.57	Coffee & Tea	2.29	Produce	1.22
Fish	0.28	Condiments & Sauces	3.14	Refrigerated Baked Goods	0.20
Fruit	5.42	Cookies & Crackers	2.69	Refrigerated Beverages	1.16
Other Meat	0.21	Dairy	12.86	Refrigerated Condiments	0.61
Other Produce	0.12	Drink Mixes	0.39	Refrigerated Desserts	0.10
Other Seafood	0.01	Ethnic	0.96	Refrigerated Dough	0.37
Pork	1.41	Frozen Baked Goods	0.34	Refrigerated Meals	0.98
Sausage	0.00	Frozen Beverages	0.20	Refrigerated Meats	5.16
Shellfish	0.15	Frozen Desserts	2.10	Snacks	6.35
Turkey	0.26	Frozen Fruits & Vegetables	1.41	Sports/Energy Drinks	0.46
Vegetables	3.59	Frozen Meals	5.19	SS Fruit	1.03
		Frozen Meat/Poultry/Seafood	4.06	SS Vegetables	1.01
		Frozen Snacks	0.43	Water	1.11

Note: There are 55 food categories in total. The average household expenditure shares of food categories are rounded to a multiple of 0.01 percent and these shares are summed up to 100 percent.

Table 4.3. Summary Statistics of Household Shopping Trip Data with Different Willingness to Travel

	Household Willingness to Travel for Shopping				
	10 km	12 km	16 km	18 km	20 km
Panel A. The Household Travel Distance to Shopping Stores					
Mean	6.49	7.73	10.23	11.44	12.63
SD	2.47	2.94	3.90	4.39	4.88
Minimum	0.31	0.31	0.31	0.31	0.31
10 Percent Quantile	2.74	3.35	4.44	4.99	5.49
25 Percent Quantile	4.68	5.62	7.50	8.22	9.03
50 Percent Quantile	6.95	8.16	10.84	12.06	13.24
75 Percent Quantile	8.59	10.19	13.52	15.12	16.79
90 Percent Quantile	9.43	11.32	15.01	16.87	18.73
Maximum	10.00	12.00	16.00	18.00	20.00
Sample Size	112,932	155,192	257,691	313,320	372,067
Panel B. The Unit Price of Household Shopping Basket					
Mean	4.11	4.12	4.12	4.13	4.13
SD	1.31	1.32	1.32	1.34	1.34
Minimum	0.61	0.61	0.61	0.61	0.61
10 Percent Quantile	3.24	3.24	3.24	3.24	3.24
25 Percent Quantile	3.52	3.53	3.53	3.53	3.53
50 Percent Quantile	3.86	3.87	3.87	3.87	3.87
75 Percent Quantile	4.28	4.29	4.29	4.30	4.30
90 Percent Quantile	4.94	4.95	4.96	4.98	4.98
Maximum	30.50	31.22	31.22	31.22	36.78
Sample Size	112,932	155,192	257,691	313,320	372,067
Panel C. The Size of Household Choice Set of Shopping Stores					
Mean	11.31	15.55	25.76	31.38	37.29
SD	4.19	5.38	8.23	9.81	11.51
Minimum	1	1	1	1	2
10 Percent Quantile	6	8	14.9	18	22
25 Percent Quantile	8	12	21	26	30
50 Percent Quantile	11	16	26	33	39
75 Percent Quantile	15	19	32	39	47
90 Percent Quantile	17	22	36	43	51
Maximum	23	28	42	51	58
Sample Size	1,280	1,280	1,280	1,280	1,280

Table 4.4. Summary Statistics of Cities Overlapped with Local Markets

County	Number of County Households	Number of Cities	Number of City Households		City Household Density (Households per km^2)	
			Mean	SD	Mean	SD
Anoka, MN	114,130	16	7,133	7,205	356	335
Carver, MN	28,439	9	3,160	3,552	113	121
Chisago, MN	2,755	1	2,755	–	52	–
Dakota, MN	153,671	19	8,088	8,781	221	198
Hennepin, MN	490,316	40	12,258	27,446	400	305
Ramsey, MN	207,327	19	10,912	25,062	461	313
Scott, MN	37,205	8	4,651	5,097	94	102
Washington, MN	86,582	25	3,463	5,494	276	407
Wright, MN	10,423	3	3,474	3,009	43	38
St. Croix, WI	1,353	1	1,353	–	16	–

Note: MN and WI are the abbreviations of Minnesota and Wisconsin, respectively. City, a subdivision of County, is a small administrative area in the United States.

Table 4.5. Summary Statistics of Sales Information of Retail Chains and Stores

Retail Chain ID	Number of Stores	Retail Chain		Store Price (Dollars per Unit)		Store Market Share (%)	
		Total Revenue (Dollars)	Market Share (%)	Mean	SD	Mean	SD
1	44	91,734,616	13.02	4.07	0.10	1.10	1.15
2	22	56,298,165	7.48	4.35	0.18	1.14	0.63
3	28	22,581,868	2.97	4.40	0.17	0.51	0.77
4	5	3,790,683	0.54	4.04	–	0.36	0.17
5	20	41,141,627	5.32	4.47	–	1.25	1.29

Note: A store price is an average price across food categories weighted by its revenue.

Table 4.6. Summary Statistics of Household Travel Distance and Shopping Basket Price

Variable	Mean	SD	Min	Max
Scenario 1. Restrict the sample to the household observations with the stores having the shortest travel distance, then to the household observations with the stores having the lowest store price.				
Travel Distance (km)	2.53	1.49	0.31	15.68
Shopping Basket Price (Dollars per Unit)	4.01	0.89	1.24	16.91
Choice (Yes = 1; No = 0)	0.39	0.49	0.00	1.00
Scenario 2. Restrict the sample to the household observations with the stores having the lowest store price, then to the household observations with the stores having the shortest travel distance.				
Travel Distance (km)	9.66	4.01	0.48	16.00
Shopping Basket Price (Dollars per Unit)	3.67	0.62	1.11	10.27
Choice (Yes = 1; No = 0)	0.25	0.43	0.00	1.00
Scenario 3. Subsample the household observations with the stores from the visited retail chains, and next restrict to the household observations with the stores having the shortest travel distance, then to the household observations with the stores having the lowest store price.				
Travel Distance (km)	3.96	2.58	0.31	15.86
Shopping Basket Price (Dollars per Unit)	4.01	0.86	1.48	15.50
Choice (Yes = 1; No = 0)	1.00	0.00	1.00	1.00
Scenario 4. Subsample the household observations with the stores from the visited retail chains, and next restrict the sample to the household observations with the stores having the lowest store price, then to the household observations with the stores having the shortest travel distance.				
Travel Distance (km)	9.13	4.39	0.46	15.97
Shopping Basket Price (Dollars per Unit)	3.90	0.78	1.34	15.50
Choice (Yes = 1; No = 0)	1.00	0.00	1.00	1.00

Note: This table is based on a subsample of the choices with the stores from retail chain 1 to 5 for each household. The choice is a binary variable indicating if a household visits the retail chain.

Table 4.7. Estimation Results from the Mixed Logit Model

Variables	Household Willingness to Travel for Shopping					Shop in
	10 km	12 km	16 km	18 km	20 km	Every Store
Shopping Basket Price						
Mean	–0.477***	–0.427***	–0.394***	–0.407***	–0.413***	–0.542***
	(0.079)	(0.090)	(0.069)	(0.072)	(0.071)	(0.076)
SD	0.742***	0.710***	0.763***	0.728***	0.727***	0.668***
	(0.032)	(0.032)	(0.035)	(0.032)	(0.032)	(0.023)
Interaction with Household Income						
≥ \$25k and <\$50k	0.247***	0.162*	0.099	0.093	0.082	0.096
	(0.082)	(0.097)	(0.078)	(0.079)	(0.078)	(0.080)
≥ \$50k and <\$70k	0.304***	0.237**	0.142*	0.123	0.141*	0.209**
	(0.098)	(0.100)	(0.083)	(0.086)	(0.086)	(0.087)
≥ \$70k and <\$100k	0.241**	0.262***	0.197**	0.179*	0.161*	0.285***
	(0.105)	(0.099)	(0.093)	(0.095)	(0.091)	(0.086)
≥ \$100k	0.219**	0.169*	0.166**	0.144*	0.133	0.216**
	(0.090)	(0.098)	(0.084)	(0.083)	(0.082)	(0.089)
Travel Distance						
Mean	–0.328***	–0.364***	–0.363***	–0.369***	–0.367***	–
	(0.025)	(0.022)	(0.021)	(0.02)	(0.018)	–
SD	0.614***	0.524***	0.465***	0.426***	0.405***	–
	(0.026)	(0.021)	(0.018)	(0.019)	(0.017)	–
Male	–1.032***	–0.846***	–0.924***	–1.026***	–1.061***	–0.838***
	(0.237)	(0.218)	(0.219)	(0.218)	(0.226)	(0.173)
Married	–0.706**	–0.807***	–1.069***	–1.041***	–1.026***	–0.738***
	(0.313)	(0.282)	(0.301)	(0.279)	(0.269)	(0.230)
White	–0.483*	–0.452	–0.182	–0.234	–0.223	–1.657***
	(0.254)	(0.294)	(0.224)	(0.234)	(0.239)	(0.194)
High School	0.086	0.042	0.576**	0.459	0.608**	–0.081
	(0.285)	(0.252)	(0.278)	(0.288)	(0.267)	(0.218)

Some College	0.328	0.332	0.280	0.416**	0.358*	–0.053
	(0.237)	(0.223)	(0.214)	(0.206)	(0.208)	(0.178)
Graduate	–1.031***	–0.901***	–0.499*	–0.304	–0.424	–0.876***
	(0.323)	(0.284)	(0.260)	(0.269)	(0.273)	(0.223)
Kid	0.202	0.165	–0.390	–0.327	–0.358	–0.196
	(0.396)	(0.353)	(0.330)	(0.321)	(0.316)	(0.300)
Owned House	–0.153	–0.033	–0.391*	–0.202	–0.229	–0.765***
	(0.248)	(0.230)	(0.232)	(0.210)	(0.219)	(0.177)
Fully Employed	–0.090	–0.030	0.227	0.032	0.079	–0.465***
	(0.246)	(0.214)	(0.200)	(0.208)	(0.202)	(0.160)
Partially Employed	0.113	0.106	–0.005	–0.073	–0.093	–1.013***
	(0.243)	(0.243)	(0.253)	(0.255)	(0.252)	(0.217)
Household Size = 2	0.686**	0.933***	0.891***	0.959***	0.987***	0.034
	(0.324)	(0.277)	(0.288)	(0.293)	(0.287)	(0.247)
Household Size = 3	0.663	0.703*	1.141***	1.072***	1.144***	–0.113
	(0.479)	(0.426)	(0.411)	(0.416)	(0.403)	(0.370)
Household Size = 4	0.943	0.633	1.469***	1.441***	1.574***	–0.064
	(0.622)	(0.564)	(0.467)	(0.437)	(0.423)	(0.413)
Household Size ≥ 5	–0.233	–0.121	0.246	0.504	0.503	–0.547
	(0.604)	(0.561)	(0.540)	(0.528)	(0.522)	(0.484)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Retail Chain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Households	1280	1280	1280	1280	1280	1280
Sample Size	112,932	155,192	257,691	313,320	372,067	1,142,618
Log–Likelihood	–10866.05	–11229.76	–11424.02	–11527.95	–11554.85	–12385.90

Note: The last column presents the estimates from the model in which households are willing to shop in every store in the entire market. To increase accuracy, we use 100 Halton draws in the estimation of these models. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table 4.8. Summary Statistics of Stores' Price Elasticities

	Number of Stores	Own-Price Elasticity Mean	Own-Price Elasticity SD	Cross-Price Elasticity Mean	Cross-Price Elasticity SD
Retail Chain ID					
1	44	–4.832	2.631	0.065	0.029
2	22	–2.123	0.892	0.026	0.009
3	28	–2.952	1.615	0.040	0.026
4	5	–3.778	1.724	0.048	0.014
5	20	–2.709	1.203	0.035	0.012
County					
Anoka	16	–3.121	1.748	0.042	0.024
Carver	4	–1.832	1.120	0.024	0.018
Dakota	27	–3.179	1.663	0.049	0.025
Hennepin	35	–3.843	1.971	0.045	0.024
Ramsey	18	–4.953	3.295	0.057	0.036
Scott	4	–2.031	1.098	0.026	0.011
Washington	15	–2.678	1.823	0.045	0.031

Table 4.9. Summary Statistics of Stores' Profit Margins (%)

	Number	Number of Rivals within the Same Site			Competition across Stores		Collusion within the Same Site		Collusion within the Retail Chain	
	of Stores	No	One	Two	Mean	SD	Mean	SD	Mean	SD
Retail Chain ID										
1	41	26	11	4	28.72	16.01	30.37	17.68	59.89	17.41
2	18	13	3	2	46.77	14.12	48.16	14.67	52.20	13.45
3	24	13	9	2	38.23	13.43	42.36	17.72	40.13	13.19
4	5	3	1	1	30.33	9.72	32.59	6.39	30.69	9.55
5	18	9	6	3	40.40	16.24	44.35	20.28	44.57	15.25
County										
Anoka	14	7	4	3	35.86	14.39	38.24	15.27	49.66	12.60
Carver	3	1	2	0	48.17	14.29	51.56	14.53	52.61	8.43
Dakota	22	9	10	3	33.44	13.19	36.19	14.82	59.00	23.40
Hennepin	33	27	6	0	33.05	15.76	33.37	15.75	44.12	14.57
Ramsey	18	11	4	3	28.77	16.18	30.30	16.23	46.11	17.97
Scott	3	3	0	0	45.18	14.92	45.18	14.92	52.05	21.50
Washington	13	6	4	3	53.06	13.19	63.18	19.11	55.43	12.34

Note: There are 106 stores with elastic own–price elasticities included in the subsequent analysis of stores' gross profit margins.

Table 4.10. Regression Results of Stores' Profit Margins on Number of Nearby Rivals within 5 km

Variable	Competition across Stores			Collusion within the Same Site			Collusion within the Retail Chain		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Constant	–28.271***	–28.421***	–25.804***	–45.475***	–45.731***	–40.853***	–18.228***	–18.179***	–15.845***
	(2.445)	(2.424)	(2.544)	(2.541)	(2.477)	(2.598)	(3.192)	(3.192)	(3.303)
Number of Nearby Stores	–0.024***			–0.018***			–0.016***		
	(0.003)			(0.003)			(0.004)		
Number of Nearby Stores		–0.047***			–0.058***			–0.009	
from Same Retail Chain		(0.006)			(0.006)			(0.007)	
Number of Nearby Stores		–0.018***			–0.008**			–0.018***	
from Other Retail Chains		(0.003)			(0.003)			(0.004)	
Number of Nearby Stores			–0.019***			–0.008*			–0.011**
from Retail Chain 1			(0.004)			(0.004)			(0.005)
Number of Nearby Stores			–0.037***			–0.043***			–0.024***
from Retail Chain 2			(0.007)			(0.007)			(0.009)
Number of Nearby Stores			–0.038***			–0.045***			–0.036***
from Retail Chain 3			(0.006)			(0.006)			(0.008)
Number of Nearby Stores			–0.041***			–0.041***			0.055***
from Retail Chain 4			(0.011)			(0.011)			(0.014)
Number of Nearby Stores			–0.002			0.024***			0.006
from Retail Chain 5			(0.007)			(0.008)			(0.010)
Pop. Percentage of Marriage	–0.141***	–0.138***	–0.120***	–0.221***	–0.215***	–0.182***	–0.106***	–0.107***	–0.084***
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.020)	(0.020)	(0.021)
Pop. Percentage of Education	0.206***	0.214***	0.176***	0.376***	0.391***	0.322***	0.205***	0.202***	0.178***
Less than High Schl Degree	(0.028)	(0.028)	(0.029)	(0.029)	(0.028)	(0.030)	(0.036)	(0.037)	(0.038)
Pop. Percentage of Education	0.196***	0.184***	0.187***	0.223***	0.201***	0.205***	0.053	0.057	0.036
High School Degree	(0.028)	(0.028)	(0.028)	(0.030)	(0.029)	(0.029)	(0.037)	(0.037)	(0.037)
Pop. Percentage of Education	0.104***	0.092***	0.098***	0.087***	0.066**	0.075***	0.011	0.015	0.000

Bachelor's Degree	(0.025)	(0.025)	(0.025)	(0.026)	(0.026)	(0.026)	(0.033)	(0.033)	(0.033)
Pop. Percentage of Education	0.052***	0.058***	0.053***	0.120***	0.132***	0.124***	0.004	0.002	0.013
Graduate Degree	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)	(0.016)
Pop. Percentage of HH Income	0.478***	0.484***	0.433***	0.807***	0.818***	0.722***	0.425***	0.422***	0.377***
≥ 25k and < 50k	(0.046)	(0.046)	(0.048)	(0.048)	(0.047)	(0.049)	(0.060)	(0.060)	(0.062)
Pop. Percentage of HH Income	0.030	0.030	0.027	0.053**	0.054**	0.048**	–0.097***	–0.098***	–0.096***
≥ 50k and < 75k	(0.023)	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.030)	(0.030)	(0.031)
Pop. Percentage of HH Income	0.369***	0.378***	0.325***	0.655***	0.671***	0.574***	0.337***	0.334***	0.302***
≥ 75k and < 100k	(0.040)	(0.039)	(0.041)	(0.041)	(0.040)	(0.042)	(0.052)	(0.052)	(0.054)
Pop. Percentage of HH Income	0.357***	0.362***	0.314***	0.605***	0.612***	0.524***	0.300***	0.299***	0.258***
≥ 100k	(0.038)	(0.037)	(0.039)	(0.039)	(0.038)	(0.040)	(0.049)	(0.049)	(0.051)
Retail Chain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272
Adjusted R^2	0.647	0.654	0.652	0.703	0.718	0.717	0.507	0.507	0.519
F Statistic	76.227	75.928	69.123	98.140	102.09	93.220	43.157	41.869	40.205

Note: The dependent variables are stores' profit margins computed under the alternative pricing scenario. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Figure 4.1. The Map of the Minneapolis and St. Paul Metropolitan Statistical Area with Sampled Stores and Households

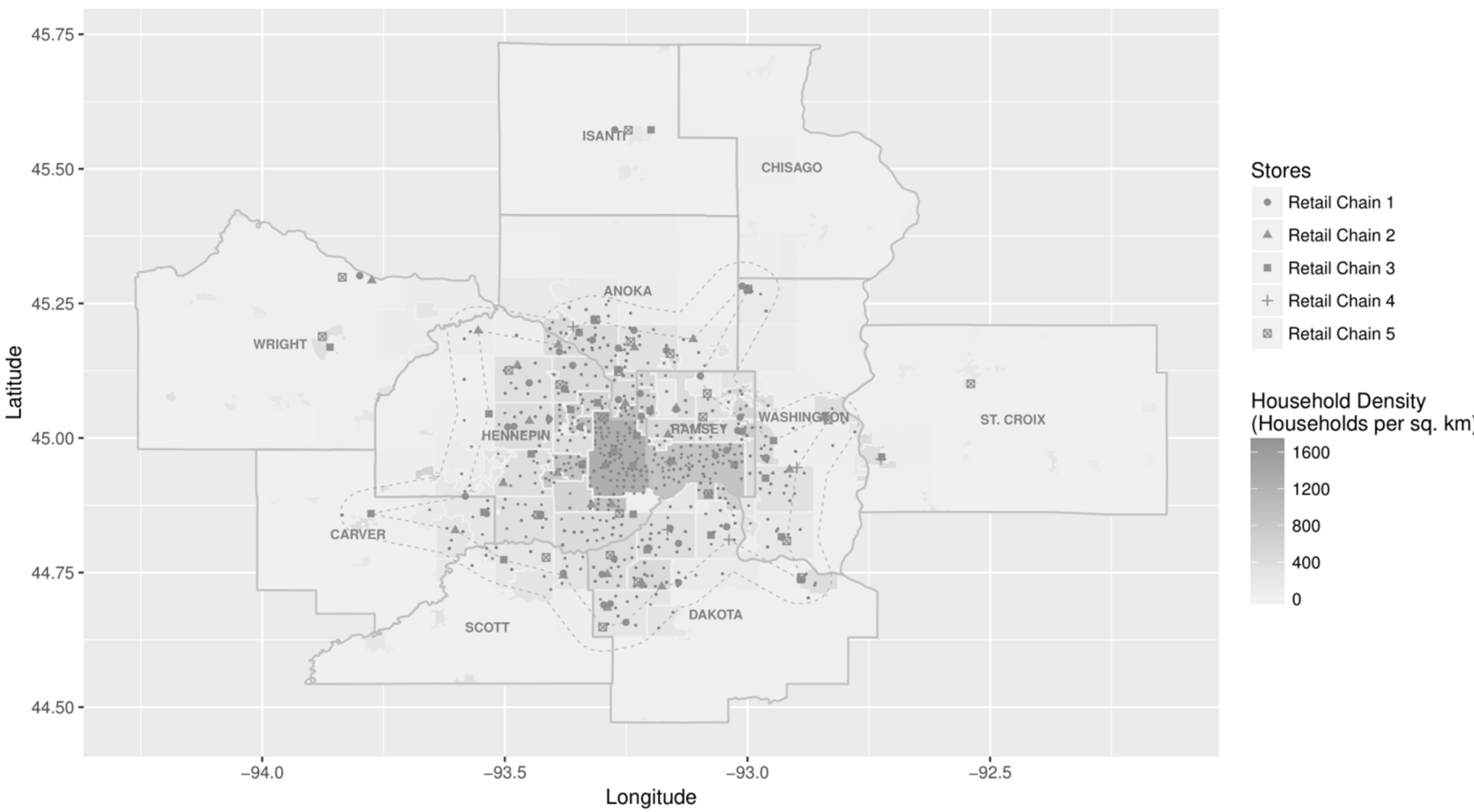


Figure 4.2. Household Willingness to Pay for Travel Distance (Dollar per km)

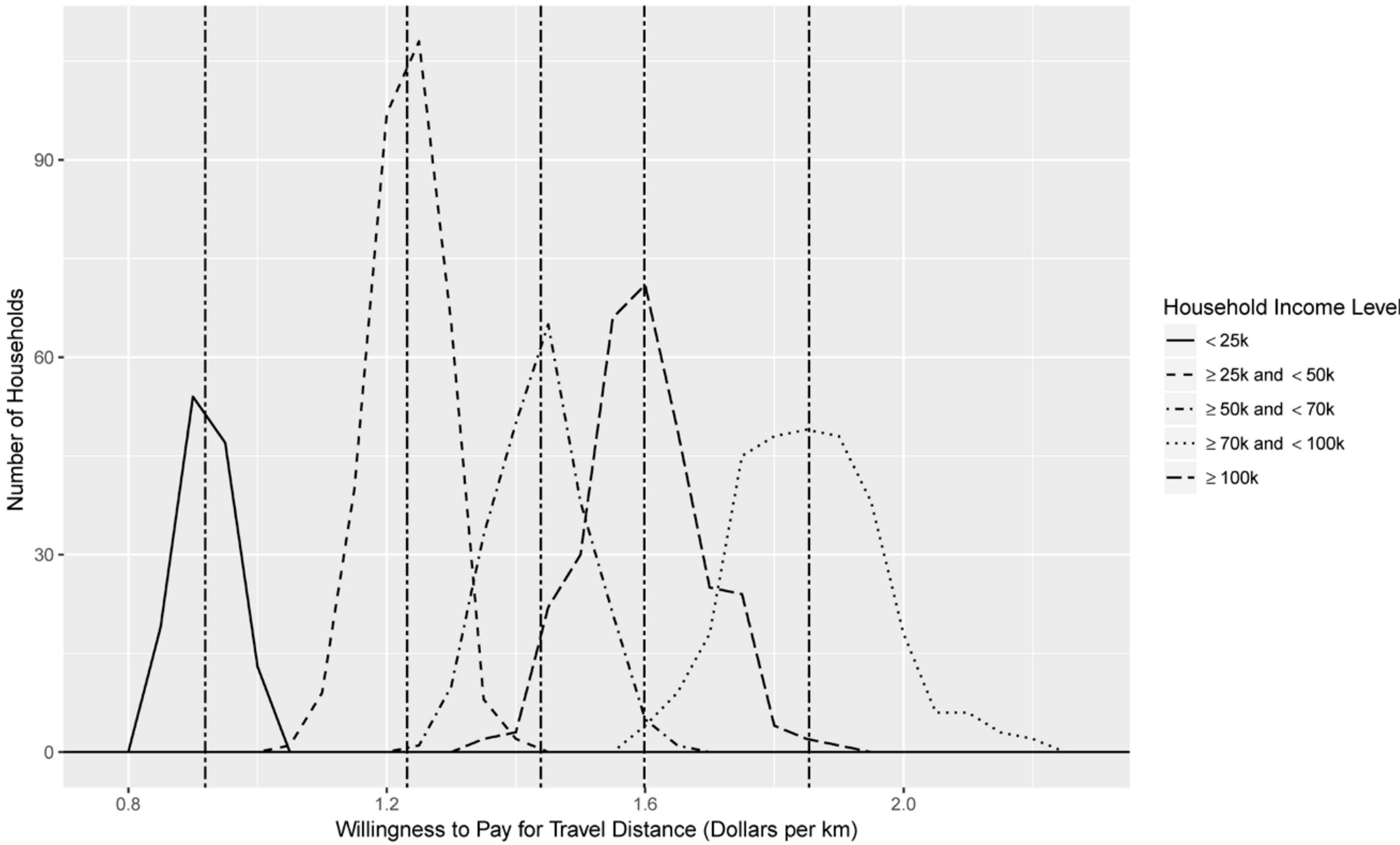
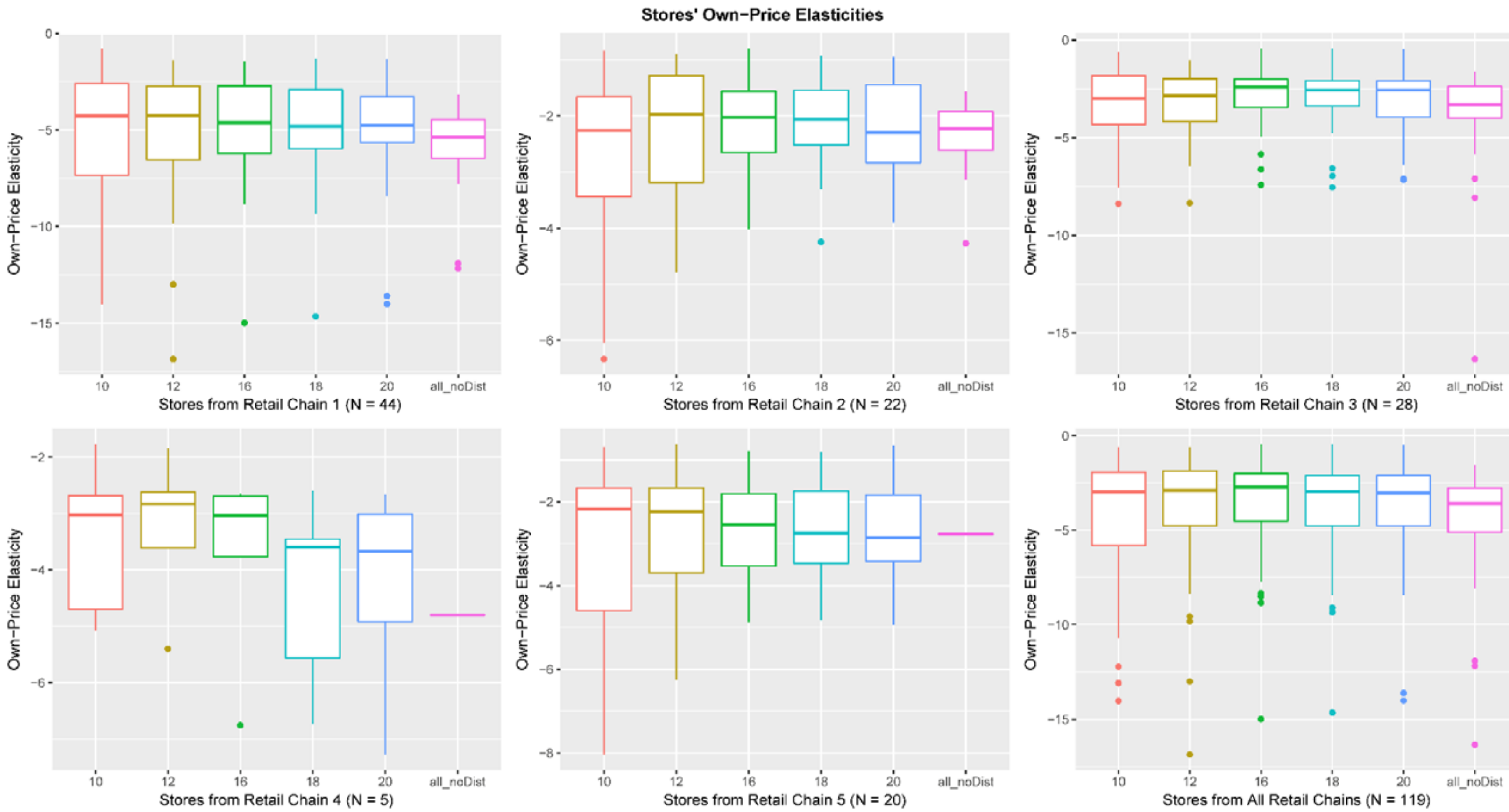
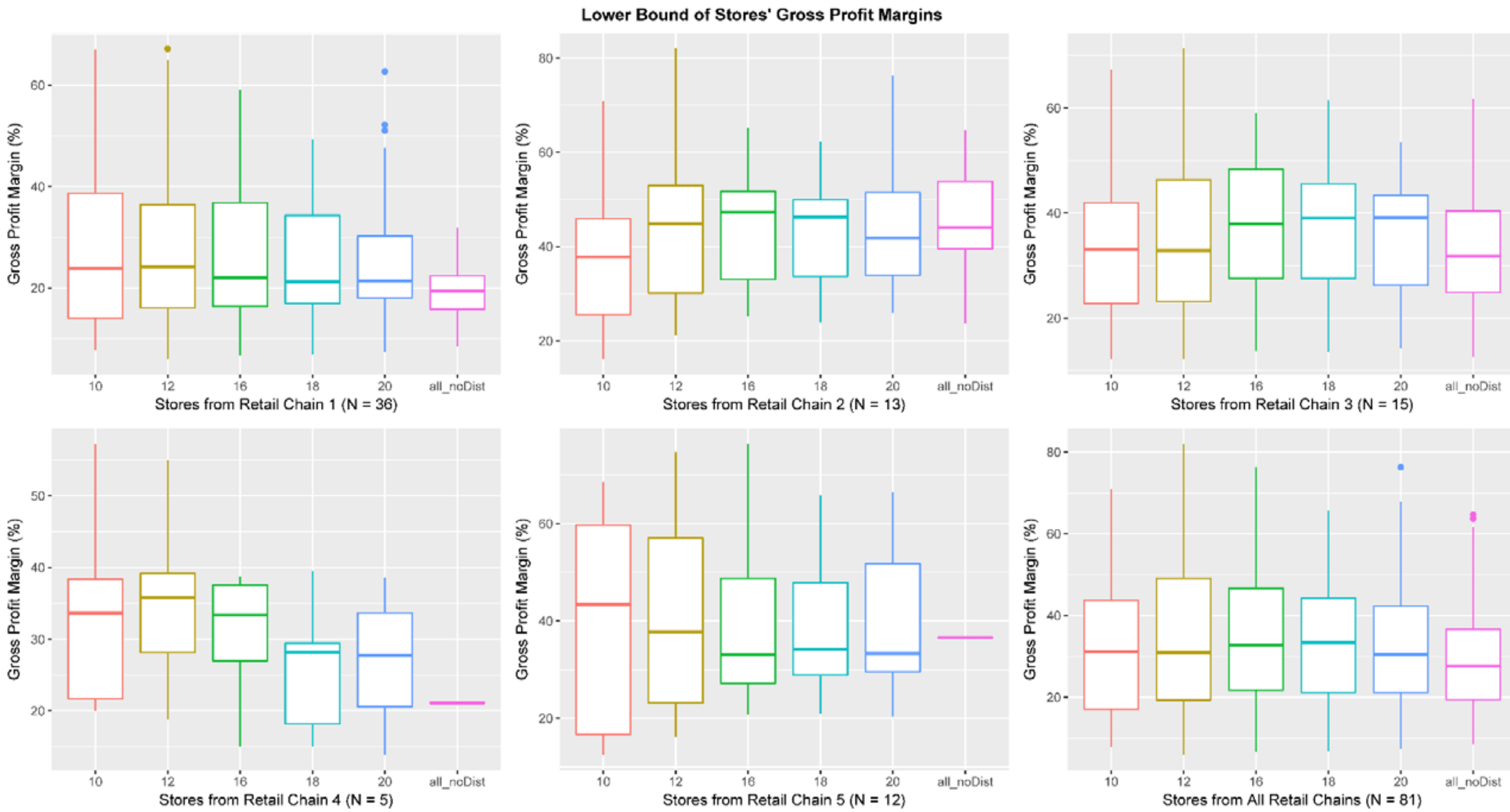


Figure 4.3. Stores' Own-Price Elasticities



Note: There are 119 stores from five retail chains in the data. The horizontal axis presents the labels of six scenarios, where the numbers represent the five scenarios in which households are willing to travel a certain distance for shopping and "all_noDist" represents the scenario in which households are willing to shop in every store in the study area regardless of travel distance. The numbers of stores are presented in the paratheses.

Figure 4.4. Stores' Gross Profit Margins (%)



Note: A collection of 81 out of 119 stores have elastic own-price elasticities in all six scenarios of household WTTFS. Stores' gross profit margins are computed in the scenario in which stores compete against each other. The horizontal axis presents the labels of six scenarios, where the numbers represent the five scenarios in which households are willing to travel a certain distance for shopping and "all_noDist" represents the scenario in which households are willing to shop in every store in the study area regardless of travel distance. The numbers of stores are presented in the paratheses.

Chapter 5. Conclusion

This dissertation comprises three essays on research questions in consumer demand for foods and consumer choices of stores. In each essay, we use a structural model of demand to elicit consumer preferences and identify product relationships via the estimates of price and income elasticities. In turn, these estimated elasticities are used to analyze the changes in welfare, policy outcomes, and firms' competition. In particular, these essays provide insights into welfare changes in response to trade restrictions, economic benefits from demand-enhancing agricultural R&D programs, and food retail competition in local markets and discuss implications for public and business policies.

The first essay focuses on the increased demand for "The New Super Grain"—Ethiopian teff—in the cereal market. Over the past decade, consumer demand for Ethiopian teff has experienced a considerable increase in Europe and North America, resulting from a growing awareness of its value in nutrition. With concerns about domestic food security, the Ethiopia government imposed an export ban on teff in 2006. This restriction was eventually released in 2015. As a consequence, teff prices increased by 10 percent in 2016 and 2017. At the same time, the government had been implementing a wheat import program since 2008 and distributed the imported wheat at subsidized prices. This chapter first investigates how much Ethiopian households suffer from the increasing teff prices, then evaluates the extent to which food aid programs could mitigate the loss of consumer welfare.

To improve our understanding of the cereal market and welfare change in Ethiopia, we estimate a two-stage food demand system by using the aggregated consumption data from the two waves of Ethiopia Socioeconomic Survey 2013-2014 and 2015-2016. The results show that among all primary cereals, teff is the most inelastic one. It implies the essential role of teff in the cereal consumption in Ethiopia. Specifically, an increase of 10 percent in teff prices needs a monetary compensation of 7.82 Birrs, which is about 3.85 percent of the weekly expenditure on foods for an average household. We also find that the policy offering wheat as food aid is a better alternative compared to the food aid policies offering other less expensive cereals, such as barley, maize, and sorghum. This finding provides empirical support for the ongoing wheat import program in Ethiopia.

The second essay focuses on the introduction of new agricultural products. As the food and agricultural market become more consumer-oriented, large numbers of demand-enhancing products have been introduced in the United States to serve consumers' heterogeneous tastes and their increasing expectations of food quality. But little is known about the economic benefits of these products. Therefore, this essay examines the social benefits from the introduction of Honeycrisp apples.

In particular, we estimate structural models of consumer demand and retailer competition using store scanner data covering 61 cities across the United States from 2009 to 2015. With the estimated demand and retailer competition model, we conduct a counterfactual analysis by simulating a scenario where Honeycrisp apples are removed from the market. We find that the introduction of Honeycrisp apples increases consumer welfare by 3.14 cents per pound, more than 90 percent of which is from the increased number of total apple varieties and the rest from the decline in prices of competing apple varieties. The extent of the decline is positively associated with the market share of Honeycrisp apples. Moreover, the simulation results show that the introduction of Honeycrisp apples has increased the total sales quantity by 8.03 percent and the total sales revenue by 21.25 percent over the study period. In addition, the success of Honeycrisp is deeply rooted in the agricultural R&D initiatives in the public university system. To provide some implications for agricultural R&D programs in the near future, we extrapolate the results to the entire U.S. apple market. We find the introduction of Honeycrisp apples has increased total consumer welfare by about 940 million dollars between 2009 and 2015, which is approximately 20 percent of the annual average spending on public food and agricultural R&D.

The third essay focuses on the food retail competition in local markets. Markets are defined by two dimensions—differentiated products and geographic areas. Stores in the food retail industry compete against their nearby rivals by providing differentiated products in neighborhood areas.

Using the 2016 IRI household and retail scanner data from a metropolitan area, we investigate the household choice of stores using a revised mixed logit model, in which we account for the heterogeneity in households' choice sets of stores, shopping baskets, and travel distances. The results suggest the importance of the investigation of store competition in local markets. We find stores' own-price elasticities and gross profit margins could be over- and under-estimated, respectively, if households are assumed to shop in every store in the market regardless of travel distance. We also find households are reluctant to travel a long distance for their grocery shopping. Instead, they would like to visit closer stores and pay for their shopping baskets at higher prices. The price coefficient implies that households would be less responsive to the change in the prices of their shopping baskets, as their incomes increase. Moreover, the average own-price elasticity (gross profit margin) over stores in a county with a larger population and higher store density is greater (lower) than the average in a county with a smaller population and lower store density. In addition, the gross profit margin of a store is negatively associated with the number of nearby rivals, regardless of alternative retail pricing strategies. Depending on different pricing strategies, one increase in the number of nearby rivals within 5 kilometers from a store is associated with a decrease of 1.6 to 2.4 percent in the store's gross profit margin.

References

Abbott, Philip C. 2012. “Export Restrictions as Stabilization Responses to Food Crisis.” *American Journal of Agricultural Economics* 94 (2): 428–34.

Alston, Julian M, Philip G Pardey, Jennifer S James, and Matthew A Andersen. 2009. “The Economics of Agricultural R&D.” *Annual Review of Resource Economics* 1: 537–66.

Assefa, Berihu. 2015. “Analysis of Price Incentives for Teff in Ethiopia for the Time Period 2005–2012.” Monitoring African Food and Agricultural Policies Project. The Food and Agriculture Organization of The United Nations. http://www.fao.org/in-action/mafap/resources/detail/en/c/394286/.

Bellemare, Marc F., Johanna Fajardo-Gonzalez, and Seth R. Gitter. 2016. “Foods and Fads-The Welfare Impacts of Rising Quinoa Prices in Peru.” Working Paper 2016–06. Towson University.

Berry, Steven. 1994. “Estimating Discrete-Choice Models of Product Differentiation.” *The RAND Journal of Economics*, 242–262.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica*, 841–890.

Bhat, Chandra R. 2001. “Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model.” *Transportation Research Part B: Methodological* 35 (7): 677–693.

Bilgic, Abdulbaki, and Steven T Yen. 2013. “Household Food Demand in Turkey: A Two-Step Demand System Approach.” *Food Policy* 43: 267–277.

Bonanno, Alessandro, and Rigoberto A Lopez. 2009. “Competition Effects of Supermarket Services.” *American Journal of Agricultural Economics* 91 (3): 555–568.

Bonnet, Céline, and Timothy J. Richards. 2016. “Models of Consumer Demand for Differentiated Products.” 16–741. TSE Working Papers. Toulouse School of Economics.

Brester, Gary W, Pascale Lhermite, Barry K Goodwin, and Melvin C Hunt. 1993. “Quantifying the Effects of New Product Development: The Case of Low-Fat Ground Beef.” *Journal of Agricultural and Resource Economics*, 239–250.

Carew, Richard, Wojciech J Florkowski, and Elwin G Smith. 2012. “Hedonic Analysis of Apple Attributes in Metropolitan Markets of Western Canada.” *Agribusiness* 28 (3): 293–309.

Central Statistical Agency. 2016. “Report on Area and Production of Major Crops.” Statistical Bulletin. Agricultural Sample Survey 2015/2016. Addis Ababa, Ethiopia: Central Statistical Agency.

Charles, Dan. 2017. "Washington Apple Growers Sink Their Teeth Into The New Cosmic Crisp." NPR. May 3, 2017. http://www.npr.org/sections/thesalt/2017/05/03/525421226/washington-apple-growers-sink-their-teeth-into-the-new-cosmic-crisp.

Chaudhuri, Shubham, Pinelopi K. Goldberg, and Panle Jia. 2006. "Estimating the Effects of Global Patent Protection in Pharmaceuticals: A Case Study of Quinolones in India." *American Economic Review* 96 (5): 1477–1514. https://doi.org/10.1257/aer.96.5.1477.

Cleeren, Kathleen, Frank Verboven, Marnik G Dekimpe, and Katrijn Gielens. 2010. "Intra-and Interformat Competition among Discounters and Supermarkets." *Marketing Science* 29 (3): 456–473.

Cotterill, Ronald W. 1986. "Market Power in the Retail Food Industry: Evidence from Vermont." *The Review of Economics and Statistics*, 379–386.

Cotterill, Ronald W. 1999. "Market Power and the Demsetz Quality Critique: An Evaluation for Food Retailing." *Agribusiness* 15 (1): 101–118.

Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *American Economic Review* 70 (3): 312–326.

Demeke, M., and F. Di Marcantonio. 2013a. "Analysis of Incentives and Disincentives for Teff in Ethiopia." Technical Notes. Monitoring African Food and Agricultural Policies Project. The Food and Agriculture Organization of The United Nations. http://www.fao.org/in-action/mafap/resources/detail/en/c/394278/.

———. 2013b. "Analysis of Incentives and Disincentives for Wheat in Ethiopia." Technical Notes. Monitoring African Food and Agricultural Policies Project. The Food and Agriculture Organization of The United Nations. http://www.fao.org/in-action/mafap/resources/detail/en/c/394278/.

Diewert, W Erwin. 2009. "Similarity Indexes and Criteria for Spatial Linking." In *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, edited by D.S. Prasada Rao, 183–216. Cheltenham UK: Edward Elgar.

Ellickson, Paul B. 2013. "Supermarkets as a Natural Oligopoly." *Economic Inquiry* 51 (2): 1142–1154.

Ethiopian Grain Trade Enterprise. 2017. "Market Statistics." http://www.egte-ethiopia.com/en/2014-04-07-05-30-44/market-statistics.html.

Foltz, Jeremy, Bradford Barham, and Kwansoo Kim. 2000. "Universities and Agricultural Biotechnology Patent Production." *Agribusiness* 16 (1): 82–95.

Gallardo, R Karina, Diem Nguyen, Vicki McCracken, Chengyan Yue, James Luby, and James R McFerson. 2012. “An Investigation of Trait Prioritization in Rosaceous Fruit Breeding Programs.” *HortScience* 47 (6): 771–776.

Götz, Linde, Thomas Glauben, and Bernhard Brümmer. 2013. “Wheat Export Restrictions and Domestic Market Effects in Russia and Ukraine during the Food Crisis.” *Food Policy* 38: 214–226.

Hausman, Jerry. 1996. “Valuation of New Goods under Perfect and Imperfect Competition.” In *The Economics of New Goods*, 207–248. University of Chicago Press.

———. 1999. “Cellular Telephone, New Products, and the CPI.” *Journal of Business & Economic Statistics* 17 (2): 188–194.

Hausman, Jerry, and Ephraim Leibtag. 2007. “Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart.” *Journal of Applied Econometrics* 22 (7): 1157–1177.

Hausman, Jerry, and Gregory Leonard. 2002. “The Competitive Effects of a New Product Introduction: A Case Study.” *The Journal of Industrial Economics* 50 (3): 237–263.

Hausman, Jerry, Gregory Leonard, and Douglas Zona. 1994. “Competitive Analysis with Differentiated Products.” *Annales d’Economie et de Statistique*, 159–180.

Ivanic, Maros, and Will Martin. 2008. “Implications of Higher Global Food Prices for Poverty in Low-Income Countries.” *Agricultural Economics* 39 (s1): 405–416.

Jeffrey, James. 2015. “Will Ethiopia’s Teff Be the next ‘Super Grain’?” *BBC News*, April 2, 2015, sec. Business. http://www.bbc.com/news/business-32128441.

Karp, David. 2015. “Beyond the Honeycrisp Apple.” *The New York Times*, November 3, 2015, sec. Food. https://www.nytimes.com/2015/11/04/dining/beyond-the-honeycrisp-apple.html.

Kim, Donghun. 2004. “Estimation of the Effects of New Brands on Incumbents’ Profits and Consumer Welfare: The US Processed Cheese Market Case.” *Review of Industrial Organization* 25 (3): 275–293.

Kyureghian, Gayaneh, Rodolfo M Nayga Jr, and Suparna Bhattacharya. 2012. “The Effect of Food Store Access and Income on Household Purchases of Fruits and Vegetables: A Mixed Effects Analysis.” *Applied Economic Perspectives and Policy* 35 (1): 69–88.

Martin, Will, and Kym Anderson. 2012. “Export Restrictions and Price Insulation During Commodity Price Booms.” *American Journal of Agricultural Economics* 94 (2): 422–27.

McCluskey, Jill J, Brady P Horn, Catherine A Durham, Ron C Mittelhammer, and Ying Hu. 2013. “Valuation of Internal Quality Characteristics across Apple Cultivars.” *Agribusiness* 29 (2): 228–241.

McCluskey, Jill J, Ron C Mittelhammer, Anna B Marin, and Khaliela S Wright. 2007. “Effect of Quality Characteristics on Consumers’ Willingness to Pay for Gala Apples.” *Canadian Journal of Agricultural Economics* 55 (2): 217–231.

McFadden, Daniel. 1981. “Econometric Models of Probabilistic Choice.” In *Structural Analysis of Discrete Data and Econometric Applications*, edited by Charles Manski and Daniel McFadden. Cambridge: The MIT Press.

Mottaleb, Khondoker A., and Dil Bahadur Rahut. 2018. “Household Production and Consumption Patterns of Teff in Ethiopia.” *Agribusiness*, January. https://doi.org/10.1002/agr.21550.

Muth, Mary K., Megan Sweitzer, Derick Brown, Kristen Capogrossi, Shawn Karns, David Levin, Abigail Okrent, Peter Siegel, and Chen Zhen. 2016. “Understanding IRI Household-Based and Store-Based Scanner Data.” Technical Bulletin 1942. Washington, DC: United States Department of Agriculture, Economic Research Service. https://www.ers.usda.gov/publications/pub-details/?pubid=47636.

Nevo, Aviv. 2000a. “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Models of Demand.” *Journal of Economics & Management Strategy* 9 (4): 513–548.

———. 2000b. “Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry.” *The RAND Journal of Economics*, 395–421.

———. 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” *Econometrica* 69 (2): 307–342.

———. 2003. “New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems.” *Review of Economics and Statistics* 85 (2): 266–275.

———. 2011. “Empirical Models of Consumer Behavior.” *Annual Review of Economics* 3 (1): 51–75.

Nurse, Earl. 2015. “The Ethiopian Superfood That Used to Be Banned.” *CNN Marketplace Africa*. http://www.cnn.com/2015/12/18/africa/ethiopian-superfood-teff/index.html.

O’Connor, Anahad. 2016. “Is Teff the New Super Grain?” *The New York Times*, August 16, 2016, sec. Eat Well. https://well.blogs.nytimes.com/2016/08/16/is-teff-the-new-super-grain/.

Pardey, Philip G, Connie Chan-Kang, Jason M Beddow, and Steven P Dehmer. 2015. “Long-Run and Global R&D Funding Trajectories: The US Farm Bill in a Changing Context.” *American Journal of Agricultural Economics* 97 (5): 1312–23.

Pardey, Philip G, Connie Chan-Kang, Steven P Dehmer, and Jason M Beddow. 2016. "Agricultural R&D Is on the Move." *Nature News* 537 (7620): 301.

Perez, Agnes, Biing-Hwan Lin, and Jane Allshouse. 2001. "Demographic Profile of Apple Consumption in the United States." FTS-292. Fruit and Tree Nuts Situation and Outlook. Washington, DC: Economic Research Service, United States Department of Agriculture.

Petrin, Amil. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy* 110 (4): 705–729.

Petrin, Amil, and Kenneth Train. 2010. "A Control Function Approach to Endogeneity in Consumer Choice Models." *Journal of Marketing Research* 47 (1): 3–13.

Pofahl, Geoffrey M, and Timothy J Richards. 2009. "Valuation of New Products in Attribute Space." *American Journal of Agricultural Economics* 91 (2): 402–415.

Porteous, Obie. 2017. "Empirical Effects of Short-Term Export Bans: The Case of African Maize." *Food Policy* 71: 17–26.

Pray, Carl E, and Keith O Fuglie. 2015. "Agricultural Research by the Private Sector." *Annual Review of Resource Economics* 7 (1): 399–424.

Provost, Claire, and Elissa Jobson. 2014. "Move over Quinoa, Ethiopia's Teff Poised to Be next Big Super Grain." *The Guardian*, January 23, 2014, sec. Global development. http://www.theguardian.com/global-development/2014/jan/23/quinoa-ethiopia-teff-super-grain.

Rashid, Shahidur. 2010. "Staple Food Prices in Ethiopia." In *Common Market for Eastern and Southern Africa Policy Seminar under the African Agricultural Marketing Project*. Maputo, Mozambique.

Rickard, Bradley J, Todd M Schmit, Miguel I Gómez, and Hao Lu. 2013. "Developing Brands for Patented Fruit Varieties: Does the Name Matter?" *Agribusiness* 29 (3): 259–272.

Santos, Adella, Nancy McGuckin, Hikari Yukiko Nakamoto, Danielle Gray, and Susan Liss. 2011. "Summary of Travel Trends: 2009 National Household Travel Survey." Trends in travel behavior, 1969-2009 FHWA-PL-ll-022. U.S. Department of Transportation Federal Highway Administration. Available at http://nhts.ornl.gov/publications.shtml.

Secorun, Laura. 2016. "Teff Could Be the next Quinoa as Ethiopia Boosts Exports." *The Guardian*, October 14, 2016, sec. Guardian Sustainable Business. http://www.theguardian.com/sustainable-business/2016/oct/14/teff-quinoa-ethiopia-boosts-exports-food-africa.

Shonkwiler, J Scott, and Steven T Yen. 1999. “Two-Step Estimation of a Censored System of Equations.” *American Journal of Agricultural Economics* 81 (4): 972–982.

Smith, Howard. 2004. “Supermarket Choice and Supermarket Competition in Market Equilibrium.” *The Review of Economic Studies* 71 (1): 235–263.

Solgaard, Hans S, and Torben Hansen. 2003. “A Hierarchical Bayes Model of Choice between Supermarket Formats.” *Journal of Retailing and Consumer Services* 10 (3): 169–180.

Stevens, Andrew. 2015. “Quinoa Quandary: Cultural Tastes and Nutrition in Peru.” Working Paper. Department of Agricultural and Resource Economics, University of California - Berkeley.

Tefera, Nigussie, Mulat Demeke, and Shahidur Rashid. 2012. “Welfare Impacts of Rising Food Prices in Rural Ethiopia: A Quadratic Almost Ideal Demand System Approach.” In *International Association of Agricultural Economists Triennial Conference*. Foz do Iguaçu, Brazil.

Turolla, Stéphane. 2016. “Spatial Competition in the French Supermarket Industry.” *Annals of Economics and Statistics*, no. 121–122: 213–59.

Unnevehr, Laurian. 1986. “Consumer Demand for Rice Grain Quality and Returns to Research for Quality Improvement in Southeast Asia.” *American Journal of Agricultural Economics* 68 (3): 634–641.

Unnevehr, Laurian, James Eales, Helen Jensen, Jayson Lusk, Jill McCluskey, and Jean Kinsey. 2010. “Food and Consumer Economics.” *American Journal of Agricultural Economics* 92 (2): 506–521.

U.S. Census Bureau. 2015. “Retail Trade: Geographic Area Series: Summary Statistics for the U.S., States, Metro Areas, Counties, and Places: 2012.” Available at American FactFinder https://factfinder.census.gov.

U.S. Department of Health and Human Services. 2017. “Healthy Food Financing Initiative.” Office of Community Services | ACF. June 14, 2017. https://www.acf.hhs.gov/ocs/programs/community-economic-development/healthy-food-financing.

USDA. 2008. “Ethiopia 2008 Crop Assessment Travel Report.” Commodity Intelligence Report. Washington, DC: Foreign Agricultural Service.

———. 2016a. “Food Availability (Per Capita) Data System.” United States Department of Agriculture, Economic Research Service. https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/food-availability-per-capita-data-system/.

———. 2016b. “Fruit and Tree Nut Yearbook: Dataset.” United States Department of Agriculture, Economic Research Service. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1377.

———. 2017a. “New Products.” April 5, 2017. https://www.ers.usda.gov/topics/food-markets-prices/processing-marketing/new-products.aspx.

USDA, Center for Nutrition Policy and Promotion. 2017b. “USDA Food Plans: Cost of Food Report for DECEMBER 2016.” USDA Food Plans: Cost of Food. United States Department of Agriculture, Center for Nutrition Policy and Promotion. https://www.cnpp.usda.gov/sites/default/files/CostofFoodDec2016.pdf.

Villas-Boas, Sofia Berto. 2007. “Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data.” *The Review of Economic Studies* 74 (2): 625–652.

Volle, Pierre. 2001. “The Short-Term Effect of Store-Level Promotions on Store Choice, and the Moderating Role of Individual Variables.” *Journal of Business Research* 53 (2): 63–73.

Volpe, Richard. 2011. “Evaluating the Performance of US Supermarkets: Pricing Strategies, Competition from Hypermarkets, and Private Labels.” *Journal of Agricultural and Resource Economics*, 488–503.

Vu, Linh, and Paul Glewwe. 2011. “Impacts of Rising Food Prices on Poverty and Welfare in Vietnam.” *Journal of Agricultural and Resource Economics*, 14–27.

Yue, Chengyan, R Karina Gallardo, James Luby, Alicia Rihn, James R McFerson, Vicki McCracken, David Bedford, et al. 2013. “An Investigation of US Apple Producers’ Trait Prioritization—Evidence from Audience Surveys.” *HortScience* 48 (11): 1378–1384.

Yue, Chengyan, and Cindy Tong. 2011. “Consumer Preferences and Willingness to Pay for Existing and New Apple Varieties: Evidence from Apple Tasting Choice Experiments.” *HortTechnology* 21 (3): 376–383.

Appendix A The Welfare Impacts of Increased Demand for The "New Super Grain"

Table A 1. Definitions of Food Segments

Segment	Items
Cereals	Teff, Wheat, Barley, Maize, and Sorghum
Pulses and Root Crops	Horse Beans, Chick Pea, Field Pea, Lentils, Haricot Beans, Potato, and Kocho
Fruit and Vegetables	Onion and Banana
Animal Products	Meat (Including Goat And Mutton Meat, Beef, Poultry, and Fish), Milk, Cheese, and Eggs

Table A 2. First–Step Probit Estimates from the Lower-Level Demand System

	Teff	Wheat	Barley	Maize	Sorghum
Constant	0.824***	0.113	–0.195	0.230	0.347**
	(0.241)	(0.178)	(0.161)	(0.161)	(0.170)
Year 2015	0.068	0.037	0.017	0.321***	0.096
	(0.124)	(0.109)	(0.095)	(0.115)	(0.098)
Region - Afar	–0.669*	0.921**	–0.703**	0.655	–1.540***
	(0.367)	(0.364)	(0.311)	(0.418)	(0.352)
Region - Amhara	–0.293	0.089	0.075	–0.164	–0.419**
	(0.273)	(0.203)	(0.177)	(0.176)	(0.187)
Region - Oromiya	–0.416	0.419**	0.560***	0.424**	–0.268
	(0.269)	(0.208)	(0.181)	(0.195)	(0.184)
Region - Somali	–3.090***	0.620**	–0.958***	0.050	–0.772***
	(0.413)	(0.270)	(0.256)	(0.291)	(0.259)
Region – Bens. Gumuz	0.153	–0.419	–0.617*	0.156	0.331
	(0.416)	(0.320)	(0.336)	(0.516)	(0.498)
Region - SNNP	–0.399	0.086	0.416**	1.168***	–0.260
	(0.258)	(0.193)	(0.176)	(0.242)	(0.181)
Region - Gambella	–0.700*	–0.968***	–0.140	0.460	–1.389***
	(0.363)	(0.324)	(0.304)	(0.566)	(0.353)
Region - Chartered Cities	–1.278***	0.927***	–0.340*	–0.242	–0.108
	(0.303)	(0.266)	(0.200)	(0.200)	(0.220)
Town	2.059***	0.644***	0.425**	0.270	0.253
	(0.458)	(0.190)	(0.165)	(0.206)	(0.164)
Urban	2.563***	1.101***	0.979***	0.044	–0.424***
	(0.353)	(0.168)	(0.129)	(0.140)	(0.126)
Major crop – teff	1.054***				
	(0.166)				
Major crop – wheat		1.679***			
		(0.227)			
Major crop – barley			1.463***		
			(0.197)		
Major crop – maize				1.344***	
				(0.183)	
Major crop – sorghum					1.878***
					(0.188)

Note: Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table A 3. Second–Step Estimates from the Lower-Level Demand System

	Teff	Wheat	Barley	Maize	Sorghum
Constant	0.120**	0.250***	0.331***	0.575***	0.599***
	(0.052)	(0.044)	(0.051)	(0.049)	(0.069)
Year 2015	–0.009	0.034**	0.017	–0.025	–0.035*
	(0.017)	(0.014)	(0.014)	(0.016)	(0.020)
Region - Afar	0.032	0.135**	0.087**	0.174***	–0.088
	(0.062)	(0.060)	(0.036)	(0.057)	(0.102)
Region - Amhara	0.079***	–0.068***	0.000	0.064**	–0.031
	(0.030)	(0.023)	(0.023)	(0.025)	(0.047)
Region - Oromiya	0.049	–0.047**	–0.026	0.104***	–0.206***
	(0.030)	(0.022)	(0.019)	(0.023)	(0.042)
Region - Somali	–0.083	0.240***	0.253***	0.155***	–0.143**
	(0.123)	(0.056)	(0.066)	(0.047)	(0.058)
Region – Bens. Gumuz	0.001	–0.068	0.114***	0.024	–0.055
	(0.068)	(0.048)	(0.039)	(0.040)	(0.072)
Region - SNNP	–0.007	–0.022	–0.003	0.133***	–0.229***
	(0.028)	(0.025)	(0.022)	(0.026)	(0.044)
Region - Gambella	0.134*	–0.010	0.002	0.270***	–0.184***
	(0.070)	(0.091)	(0.025)	(0.066)	(0.054)
Region - Chartered Cities	0.146***	–0.026	0.056**	0.045*	–0.015
	(0.029)	(0.025)	(0.024)	(0.023)	(0.045)
Town	0.228***	–0.103***	–0.126***	–0.184***	–0.138***
	(0.024)	(0.021)	(0.018)	(0.020)	(0.024)
Urban	0.353***	–0.157***	–0.174***	–0.180***	–0.189***
	(0.023)	(0.016)	(0.019)	(0.018)	(0.033)
$\ln(X_G / P_G)$	0.059***	0.031***	–0.012	–0.105***	–0.030*
	(0.015)	(0.012)	(0.017)	(0.013)	(0.017)
$\ln(P_{\text{Teff}})$	0.098***	–0.022	0.017	–0.034*	–0.024
	(0.027)	(0.020)	(0.021)	(0.018)	(0.019)
$\ln(P_{\text{Wheat}})$	–0.022	0.062**	–0.066***	0.010	0.039**
	(0.020)	(0.025)	(0.013)	(0.018)	(0.016)
$\ln(P_{\text{Barley}})$	0.017	–0.066***	0.035	0.003	–0.004
	(0.021)	(0.013)	(0.031)	(0.017)	(0.014)
$\ln(P_{\text{Maize}})$	–0.034*	0.010	0.003	0.030	–0.084***
	(0.018)	(0.018)	(0.017)	(0.031)	(0.018)
$\ln(P_{\text{Sorghum}})$	–0.024	0.039**	–0.004	–0.084***	0.055*
	(0.019)	(0.016)	(0.014)	(0.018)	(0.030)
Inverse Mills Ratio	–0.156***	–0.155***	–0.290***	–0.187***	–0.163***
	(0.048)	(0.034)	(0.039)	(0.040)	(0.033)

Note: Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table A 4. Estimates from the Upper-Level Demand System

	Cereals	Pulses and Root Crops	Fruit and Vegetables	Animal Products
Constant	0.713***	0.172***	0.145***	–0.030
	(0.036)	(0.028)	(0.014)	(0.029)
Year 2015	0.013	0.006	0.002	–0.021**
	(0.012)	(0.010)	(0.004)	(0.010)
Region - Afar	–0.305***	–0.045	0.025*	0.325***
	(0.041)	(0.029)	(0.013)	(0.036)
Region - Amhara	–0.084***	0.070***	–0.003	0.017
	(0.022)	(0.016)	(0.006)	(0.019)
Region - Oromiya	–0.058***	0.040**	0.010*	0.007
	(0.022)	(0.016)	(0.006)	(0.019)
Region - Somali	–0.403***	0.003	0.014	0.386***
	(0.035)	(0.024)	(0.011)	(0.029)
Region – Bens. Gumuz	–0.204***	0.066**	0.046***	0.092**
	(0.041)	(0.030)	(0.011)	(0.036)
Region - SNNP	–0.192***	0.178***	0.013***	0.001
	(0.021)	(0.015)	(0.005)	(0.018)
Region - Gambella	–0.254***	–0.007	0.015	0.246***
	(0.039)	(0.028)	(0.010)	(0.034)
Region - Chartered Cities	–0.177***	0.039*	0.019**	0.119***
	(0.028)	(0.020)	(0.009)	(0.024)
Town	0.035*	–0.033**	0.030***	–0.032*
	(0.021)	(0.015)	(0.007)	(0.018)
Urban	–0.067***	–0.012	0.042***	0.037**
	(0.017)	(0.012)	(0.005)	(0.015)
$\ln(X/P)$	–0.040***	–0.007	–0.033***	0.080***
	(0.012)	(0.009)	(0.004)	(0.011)
$\ln(P_{\text{Cereals}})$	0.117***	–0.032***	–0.030***	–0.055***
	(0.013)	(0.009)	(0.008)	(0.009)
$\ln(P_{\text{Pul\&RCrp}})$	–0.032***	0.047***	–0.005	–0.011*
	(0.009)	(0.010)	(0.005)	(0.006)
$\ln(P_{\text{Fruit\&Veg}})$	–0.030***	–0.005	0.031***	0.004**
	(0.008)	(0.005)	(0.006)	(0.002)
$\ln(P_{\text{AnimalProd}})$	–0.055***	–0.011*	0.004**	0.062***
	(0.009)	(0.006)	(0.002)	(0.008)

Note: Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

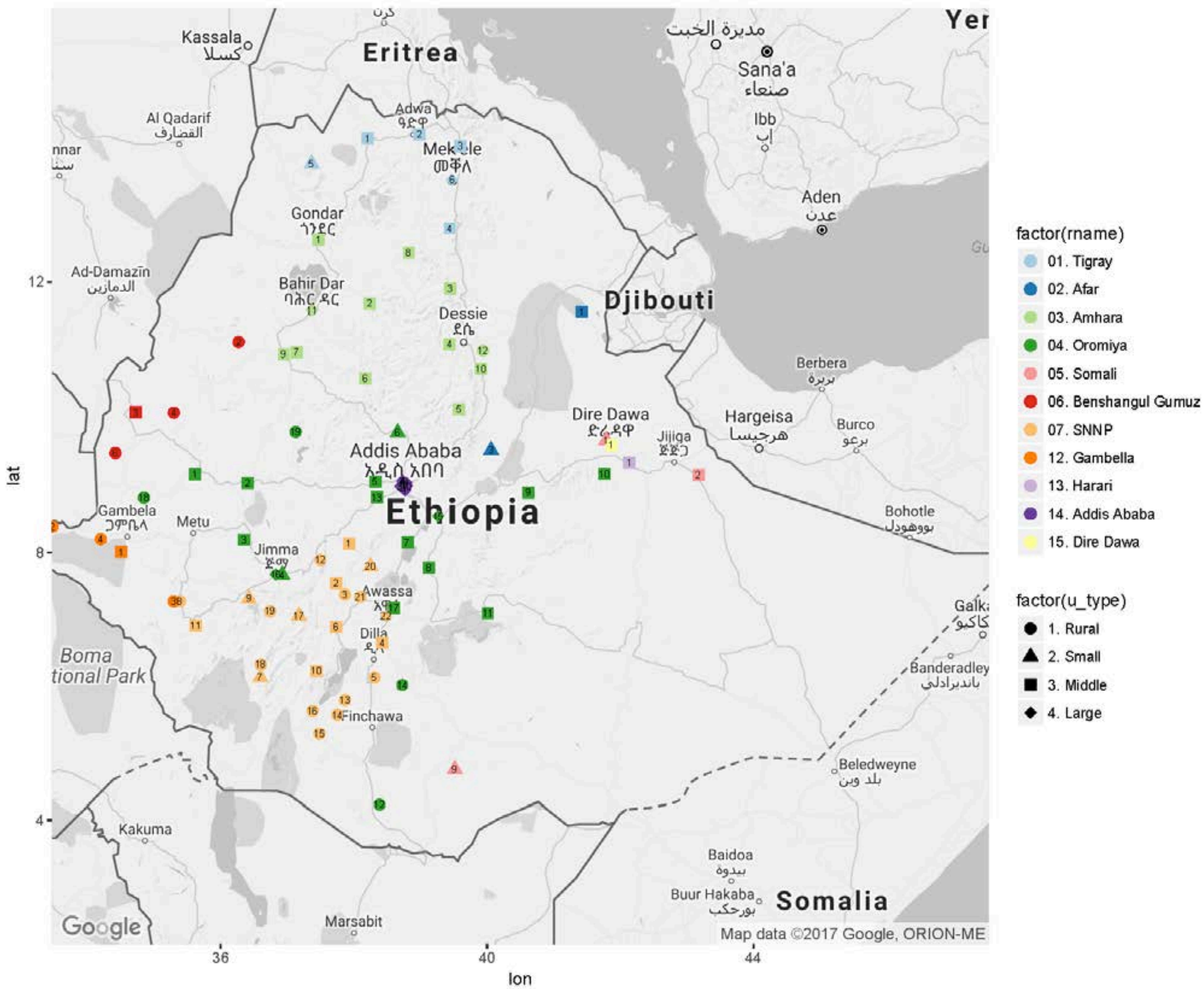
Table A 5. Policy Outcomes of Weekly Cereal Consumption as Teff Prices Increase (Kilogram)

Policy	Teff	Wheat	Barley	Maize	Sorghum	Dissimilarity from Cash Voucher
Weekly cereal consumption for an average household from the ESS 2015/2016						
	4.21	2.97	1.02	3.76	2.51	
Panel A. Teff prices increase by 5 percent						
Weekly cereal consumption for an average household simulated using the estimated demand system as teff prices increase by 5 percent with CV as 3.77 Birrs						
Food aid – cash voucher	4.00	3.07	1.19	5.94	3.23	0.00
Food aid – wheat	3.90	3.52	1.17	5.87	3.17	0.39
Food aid – barley	3.90	2.99	1.70	5.87	3.17	1.12
Food aid – maize	3.90	2.99	1.17	6.77	3.17	0.41
Food aid – sorghum	3.90	2.99	1.17	5.87	3.87	0.55
Weekly cereal consumption for an average household simulated using the estimated elasticities as teff prices increase by 5 percent with CV as 2.30 Birrs						
Food aid – cash voucher	4.13	3.01	1.03	3.74	2.52	0.00
Food aid – wheat	4.07	3.29	1.03	3.73	2.50	0.17
Food aid – barley	4.07	2.97	1.35	3.73	2.50	0.56
Food aid – maize	4.07	2.97	1.03	4.28	2.50	0.29
Food aid – sorghum	4.07	2.97	1.03	3.73	2.93	0.31
Panel B. Teff prices increase by 15 percent						
Weekly cereal consumption for an average household simulated using the estimated demand system as teff prices increase by 15 percent with CV as 12.16 Birrs						
Food aid – cash voucher	4.12	3.22	1.33	6.25	3.40	0.00
Food aid – wheat	3.81	4.67	1.25	6.00	3.20	3.24
Food aid – barley	3.81	2.97	2.97	6.00	3.20	7.73

Food aid – maize	3.81	2.97	1.25	8.92	3.20	3.39
Food aid – sorghum	3.81	2.97	1.25	6.00	5.44	4.34
Weekly cereal consumption for an average household simulated using the estimated elasticities as teff prices increase by 15 percent with CV as 6.75 Birrs						
Food aid – cash voucher	3.96	3.10	1.07	3.69	2.55	0.00
Food aid – wheat	3.80	3.93	1.05	3.68	2.49	1.27
Food aid – barley	3.80	2.99	2.00	3.68	2.49	3.81
Food aid – maize	3.80	2.99	1.05	5.29	2.49	2.21
Food aid – sorghum	3.80	2.99	1.05	3.68	3.74	2.26

Note: The first part shows the weekly cereal consumptions for an average household at the expenditure of 203 Birrs from data. Panel A and Panel B show changes in the weekly household cereal consumption as teff prices increase by 5 and 15 percent, respectively. In each panel, there are two parts, presenting the simulation results using the estimated demand system and elasticities. According to Equation 14, the last column shows the value of dissimilarity measure ($\times$ 100) between food aid policy offering cash–voucher and that providing wheat only, barley only, maize only, and sorghum only.

Figure A 1. Maps of Markets in Ethiopia



Appendix B The Welfare Impacts of New Demand-Enhancing Agricultural Products

B.1. Data

The data described in Section 4.3.4 are used for the estimation of demand. The main data are from the primary IRI InfoScan data, including weekly sales revenue and quantity from the "census" retailers at the Universal Product Codes (UPC) level. The "census" retailers are referred to those that have agreed to contribute their sales data to the IRI database. The data purchased by the USDA include only the sales data from these "census" retailers, which are an unprojected subset of the full IRI InfoScan data (Muth et al. 2016). According to agreements between the IRI and the data providing retailers, some of the InfoScan data are collected at the store level, while others are collected at the retailer market area (RMA) level. The geographic areas of the RMAs, covering several states, are self-defined and different by retailers. Therefore, the sales revenue and quantity of RMA retailers cannot be separated by (IRI) city. For a clear definition of the market, only non-RMA retailers are included in this chapter. To provide insights into the degree to which these two types of retailers have systematic differences in the context of this study, we compare the distribution of apple sales quantity from RMA and non-RMA retailers over the study period in Table B1. The table shows that Honeycrisp is sold in both types of retailers and display similar increasing trends in market share. In 2014, Honeycrisp became the 5th most popular apple in both types of retailers with average market shares of 7.3 percent and 4.6 percent in non-RMA and RMA retailers, respectively.

B.2. Apple Characteristics and Consumer Demographics

Table B2 presents apple characteristics by variety. These data are obtained from the apple variety information provided by the Washington Apple Commission. The information includes a collection of expert assessments for usage (e.g., pie stuffing, applesauce, baking, and freezing) and a measure of sweetness. In practice, we express expert assessments in binary variables, where 1 refers to "Excellent" and 0 otherwise. The variables of consumer demographics, such as age and household income, are sampled from the American Community Survey. The American Community Survey contains annual population statistics for age and household income by age at the county level. In line with the sales information, we aggregate these statistics at (IRI) city level to obtain empirical distributions of age and household income. Accordingly, we sample 1000 consumers for every market from their corresponding distributions. Table B3 describes the sample statistics for age and household income.

B.3. Retailer Groups

Apples are assumed to be differentiated by variety and by retailer. The retailers in our sample are divided into four groups based on their size: local retailers, small regional retailers, regional retailers, and nationwide retailers. Table B4 shows the distribution of retailers by size. Table B5 presents average prices and market shares of apples by variety and by retailer. The descriptive statistics show that retail prices are notably different across groups. In particular, compared to other retailers, the nationwide retailers sell all varieties but Golden Delicious at the lowest prices, while the regional retailers sell all varieties at the highest prices. Table B5 also shows that the local and small regional retailers account for the majority of Honeycrisp sales.

B.4. Terminal Market Prices

The USDA Agricultural Marketing Service (AMS) provides data on monthly average prices for apples by variety from 15 selected terminal markets across the United States. We construct the terminal market prices for 61 cities in our sample as follows. If an IRI city has a terminal market, then retailers in that city pay the prices reported in that terminal market. If an IRI city does not have a terminal market, then retailers in that city are assumed to pay the average of prices reported in terminal markets that are in the same division.[48] For example, terminal market prices in New York are assigned as prices for retailers in Buffalo, Syracuse, and Albany. If an IRI city does not have a terminal market within its division, then retailers in that division are assumed to pay the average of prices reported in terminal markets in the adjacent division. For example, retailers in Phoenix are assumed to pay an average of prices reported in Los Angeles and San Francisco. Table B6 presents the full list of IRI cities and their corresponding terminal prices.

[48] The United States Census Bureau defines four statistical regions with nine divisions for data collection and analysis.

Table B 1. Comparison of Sales Quantity from non-RMA and RMA Retailers (Percent)

	2009		2010		2011		2012		2013		2014	
Sales Quantity	Non-RMA	RMA	Non-RMA	RMA	Non-RMA	RMA	Non-RMA	RMA	Non-RMA	RMA	Non-RMA	RMA
Gala	17.70	25.88	19.24	27.48	22.01	29.30	22.72	28.83	25.45	27.79	23.84	30.03
Red Delicious	20.11	27.77	18.65	27.56	17.32	23.99	18.17	23.50	16.67	20.85	17.24	16.90
Fuji	11.99	12.80	9.69	12.23	8.50	13.66	9.13	14.62	9.66	17.33	8.25	14.38
Granny Smith	12.12	10.67	11.30	10.48	12.25	10.83	12.84	9.49	12.12	9.46	12.14	9.99
Honeycrisp	2.48	1.39	3.64	2.24	5.27	2.44	5.33	2.84	6.82	3.37	7.34	4.58
Golden Delicious	6.06	5.65	5.67	5.11	4.98	4.78	4.37	4.41	4.08	4.19	3.98	4.25
Pink Lady/Cripps Pink	1.72	0.86	1.99	1.32	2.62	1.56	3.16	1.65	3.64	1.93	3.27	2.28
Braeburn	2.97	1.84	3.02	1.58	2.41	1.32	2.06	0.96	1.97	0.84	1.54	0.65

Table B 2. Apple Characteristics by Variety

Variety	Pie	Sauce	Baking	Freezing	Sweetness
Gala	Very Good	Excellent	Very Good	Not Suggested	0.83
Red Delicious	Not Suggested	Not Suggested	Not Suggested	Not Suggested	0.33
Fuji	Very Good	Very Good	Very Good	Very Good	0.93
Granny Smith	Excellent	Excellent	Excellent	Excellent	0.08
Honeycrisp	Excellent	Excellent	Excellent	Good	0.67
Golden Delicious	Excellent	Excellent	Excellent	Excellent	0.56
Pink Lady/Cripps Pink	Excellent	Excellent	Very Good	Very Good	0.17
Braeburn	Very Good	Very Good	Very Good	Very Good	0.39

Note: The variety information is given by the Washington Apple Commission. The measure of sweetness is monotonically normalized from 0 to 1. As a result, a sweeter apple variety will have a larger measure.

Table B 3. Sample Statistics for Consumer Demographics

	Mean	SD	Min	Max
Age (Years)	49.62	17.04	18	85
Household Income ($1000)	67.49	51.30	10	200
Young Adult (25-44 Years Old)	0.36	0.48	0	1

Note: Consumer demographic variables are sampled from the American Community Survey provided by the United States Census Bureau. Young adult is defined as a binary indicator for a consumer aged between 25 and 44.

Table B 4. Distribution of Retailers by Size

	Local							Small Regional					Regional		Nationwid	
Numb. of covered IRI cities	1	2	3	4	5	6	7	10	13	15	18	19	22	28	60	61
Numb. of non-RMA retailer(s)	14	11	6	2	4	2	1	1	1	1	1	1	1	2	1	1
Composition by channel type																
Convenience	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
Defense commissary	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Dollar	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-
Drug	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	1
Grocery	14	11	5	2	3	2	1	1	-	1	1	1	1	-	-	-
Mass merchandise	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1	-

Table B 5. Sales Information by Variety and Outlet

	Local		Small Regional		Regional		Nationwide	
Variety	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Prices (Dollars per Pound)								
Gala	0.40	0.17	0.33	0.11	0.52	0.16	0.28	0.09
Red Delicious	0.33	0.14	0.34	0.12	0.45	0.16	0.32	0.12
Fuji	0.59	0.18	0.60	0.21	0.61	0.20	0.37	0.13
Granny Smith	0.45	0.16	0.38	0.11	0.53	0.17	0.27	0.08
Honeycrisp	1.14	0.36	1.11	0.26	1.15	0.32	0.64	0.34
Golden Delicious	0.53	0.18	0.51	0.18	0.67	0.15	0.59	0.29
Pink Lady/Cripps Pink	0.78	0.22	0.77	0.20	0.83	0.22	0.39	0.13
Braeburn	0.73	0.18	0.74	0.20	0.77	0.18	0.46	0.22
Market Shares (Percent)								
Gala	3.75	5.09	4.18	4.90	1.03	1.35	0.69	0.76
Red Delicious	3.58	3.75	3.11	3.55	1.01	1.37	0.33	0.33
Fuji	1.63	2.87	0.99	1.27	1.00	2.02	0.31	0.38
Granny Smith	1.95	2.04	2.27	2.50	0.81	0.97	0.43	0.44
Honeycrisp	0.95	2.04	0.69	1.42	0.46	0.88	0.46	1.02
Golden Delicious	1.10	1.54	0.90	1.12	0.24	0.28	0.15	0.33
Pink Lady/Cripps Pink	0.45	0.64	0.45	0.70	0.23	0.31	0.27	0.43
Braeburn	0.40	0.70	0.59	1.55	0.20	0.33	0.10	0.24

Note: Prices are deflated by regional price indices from the Bureau of Labor Statistics (BLS) with the base period at 1982-84.

Table B 6. IRI Cities and Terminal Markets

IRI Cities	Terminal Markets
BOS, HAS, PRO	Boston
NYC, BUF, SYR, ALB	New York
HAR	Average over Philadelphia and Pittsburgh
PHL	Philadelphia
PIT	Pittsburgh
DET, GRR	Detroit
TOL, CLE, COL	Average over Detroit and Pittsburgh
CIN, LOU	Average over St. Louis, Chicago, and Detroit
CHI, IND, MIL, GRB	Chicago
STL, KAN, WIC	St. Louis
PEO, MSP, DSM, OMA	Average over St. Louis and Chicago
BAL, RIC, ROA	Baltimore
CHL, RAL	Columbia
ATL, BIR	Atlanta
MIA, TAM, ORL, JAC	Miami
NAS, MEM, KNX	Average over St. Louis and Atlanta
DAL, NOL, HOU, SAT	Dallas
OKL, TUL, LIT	Average over St. Louis and Dallas
SLC, DEN	Average over Seattle, Los Angeles, and San Francisco
PHX, LAS	Average over Los Angeles and San Francisco
LAX, SDC	Los Angeles
SFC, SAC	San Francisco
SEA, PRT, SPK, BOI	Seattle

Note: In Figure 3.2, the IRI cities are denoted by shadowed areas with associated labels, while the terminal markets are marked by circles. The abbreviations are spelled out in the continued table.

Table B 6. IRI Cities and Terminal Markets (Continued)

Abbreviation	IRI City	Abbreviation	IRI City	Abbreviation	IRI City
ALB	Albany, NY	IND	Indianapolis, IN	PHX	Phoenix/Tucson, AZ
ATL	Atlanta, GA	JAC	Jacksonville, FL	PIT	Pittsburgh, PA
BAL	Baltimore, MD/Washington, DC	KAN	Kansas City, KS	PRO	Providence, RI
BIR	Birmingham/Montgomery, AL	KNX	Knoxville, TN	PRT	Portland, OR
BOI	Boise, ID	LAS	Las Vegas, NV	RAL	Raleigh/Greensboro, NC
BOS	Boston, MA	LAX	Los Angeles, CA	RIC	Richmond/Norfolk, VA
BUF	Buffalo/Rochester, NY	LIT	Little Rock, AR	ROA	Roanoke, VA
CHI	Chicago, IL	LOU	Louisville, KY	SAC	Sacramento, CA
CHL	Charlotte, NC	MEM	Memphis, TN	SAT	San Antonio/Corpus Christi, TX
CIN	Cincinnati/Dayton, OH	MIA	Miami/Ft Lauderdale, FL	SDC	San Diego, CA
CLE	Cleveland, OH	MIL	Milwaukee, WI	SEA	Seattle/Tacoma, WA
COL	Columbus, OH	MSP	Minneapolis/St Paul, MN	SFC	San Francisco/Oakland, CA
DAL	Dallas/Ft Worth, TX	NAS	Nashville, TN	SLC	Salt Lake City, UT
DEN	Denver, CO	NOL	New Orleans, LA/Mobile, AL	SPK	Spokane, WA
DET	Detroit, MI	NYC	New York, NY	STL	St Louis, MO
DSM	Des Moines, IA	OKL	Oklahoma City, OK	SYR	Syracuse, NY
GRB	Green Bay, WI	OMA	Omaha, NE	TAM	Tampa/St Petersburg, FL
GRR	Grand Rapids, MI	ORL	Orlando, FL	TOL	Toledo, OH
HAR	Harrisburg/Scranton, PA	PEO	Peoria/Springfield, IL	TUL	Tulsa, OK
HAS	Hartford, CT/Springfield, MA	PHL	Philadelphia, PA	WIC	Wichita, KS
HOU	Houston, TX				

Appendix C Food Retail Competition in Local Markets

Table C 1. Average Shares of Store Sales Revenue of Food Categories (%)

	Retail Chain 1		Retail Chain 2		Retail Chain 3		Retail Chain 4		Retail Chain 5	
Food Category	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Perishable Product Items										
Baked Goods	2.04	0.42	5.32	0.37	11.06	5.18	1.04	–	1.11	–
Beef	6.66	0.37	8.93	0.22	9.69	0.87	6.32	–	6.40	–
Beef/Veal	8.19	3.59							9.91	–
Chicken	3.12	0.31	5.96	0.14	7.59	0.82	2.25	–	2.45	–
Deli Cheese	6.80	0.53	5.83	0.11			8.08	–	7.67	–
Deli Meat	7.14	0.18	7.53	0.13			7.16	–	6.56	–
Deli Prepared	6.24	0.20	5.39	0.25	4.72	0.75	5.52	–	5.24	–
Fish	8.41	0.54	11.40	0.82			6.72	–	6.42	–
Fruit	2.72	0.13	3.25	0.09	2.89	0.28	2.26	–	2.24	–
Other Meat	3.73	0.35	4.43	0.26	3.13	0.18	2.78	–	2.90	–
Other Produce	20.63	3.40	21.92	1.79	49.54	7.06	8.49	–	9.66	–
Other Seafood	7.47	0.83	3.13				5.92	–	29.09	–
Pork	3.16	0.15	3.07	0.09	3.10	0.11	2.87	–	2.98	–
Sausage	4.77	0.59					2.73	–	3.66	–
Shellfish	11.50	0.91	10.28	0.62	16.00	2.18			5.68	–
Turkey	2.55	0.24	2.00	0.28	1.23	0.09	2.19	–	1.99	–
Vegetables	2.38	0.12	2.56	0.06	2.67	0.55	2.15	–	2.07	–
Processed Product Items										
Baby Food	12.58	1.67	15.56	1.01	17.11	0.90	14.82	–	15.47	–
Bakery	3.85	0.19	3.03	0.04	2.76	0.09	2.59	–	2.76	–
Baking	3.57	0.13	3.96	0.19	3.85	0.28	3.55	–	3.84	–
Breakfast	3.45	0.07	3.36	0.10	3.30	0.10	4.12	–	4.30	–
Candy	3.55	0.16	3.56	0.07	3.56	0.09	3.85	–	3.98	–
Carbonated Soft Drinks	3.66	0.10	2.97	0.04	2.88	0.11	3.74	–	4.13	–
Coffee & Tea	6.35	0.23	7.90	0.22	7.77	0.50	7.40	–	7.53	–
Condiments & Sauces	3.11	0.10	3.33	0.09	3.41	0.13	3.27	–	3.52	–
Cookies & Crackers	3.48	0.09	3.12	0.08	3.12	0.09	3.13	–	3.18	–
Dairy	3.34	0.09	3.32	0.07	3.13	0.09	3.45	–	3.77	–

Drink Mixes	3.57	0.27	4.26	0.14	4.32	0.24	3.62	–	4.02	–
Ethnic	2.93	0.08	2.84	0.05	2.76	0.06	2.66	–	2.96	–
Frozen Baked Goods	4.05	0.10	3.49	0.11	3.89	0.25	4.10	–	3.97	–
Frozen Beverages	2.08	0.12	2.56	0.09	2.57	0.11	1.63	–	1.72	–
Frozen Desserts	4.35	0.11	4.02	0.09	3.83	0.05	4.46	–	4.64	–
Frozen Fruits & Vegetables	2.89	0.14	4.65	0.28	4.79	0.47	3.24	–	3.32	–
Frozen Meals	4.29	0.13	4.13	0.07	4.04	0.09	4.34	–	4.63	–
Frozen Meat/Poultry/Seafood	7.30	0.19	6.97	0.09	6.50	0.23	6.85	–	7.01	–
Frozen Snacks	5.31	0.36	4.64	0.13	4.34	0.13	5.52	–	5.96	–
Juices	2.98	0.14	3.09	0.08	3.12	0.15	2.92	–	3.08	–
Liquor	8.66	0.27	9.42	2.28	2.90	1.89	9.38	–	15.38	–
Meals	2.43	0.15	2.25	0.07	2.12	0.09	2.28	–	2.44	–
Non–Fruit Drinks	3.74	0.35	3.97	0.25	3.86	0.54	7.40	–	5.31	–
Other Frozen	3.06	0.53	3.32	0.20	3.19	0.25			3.35	–
Other Refrigerated	2.97	0.10	2.35	0.03	2.34	0.05	2.63	–	2.77	–
Produce	3.09	0.11	3.26	0.08	3.15	0.23	3.28	–	3.31	–
Refrigerated Baked Goods	7.31	1.79	3.61	0.56	2.11	0.20	1.66	–	1.81	–
Refrigerated Beverages	4.07	0.13	3.52	0.04	3.46	0.09	3.26	–	3.59	–
Refrigerated Condiments	3.81	0.12	3.76	0.07	3.30	0.34	3.27	–	3.36	–
Refrigerated Desserts	5.87	0.55	5.43	0.37	3.86	0.14	2.83	–	2.87	–
Refrigerated Dough	3.11	0.24	2.67	0.11	2.56	0.08	2.60	–	2.68	–
Refrigerated Meals	6.00	0.54	4.45	0.22	3.97	0.43	3.49	–	4.19	–
Refrigerated Meats	5.75	0.18	5.66	0.07	4.96	0.40	4.55	–	5.29	–
Snacks	3.84	0.18	4.51	0.12	4.64	0.24	4.13	–	4.15	–
Sports/Energy Drinks	3.73	0.23	5.26	0.17	5.26	0.35	5.80	–	5.83	–
SS Fruit	2.89	0.15	3.26	0.13	3.33	0.18	2.75	–	3.07	–
SS Vegetables	1.58	0.12	1.64	0.11	1.47	0.17	1.57	–	1.66	–
Water	2.79	0.16	3.04	0.10	3.03	0.12	3.44	–	3.52	–

Note: Retail chain 1, 2, and 3 report sales information at the store level, while retail chain 4 and 5 report at the Retail Market Area (RMA) level. The RMA is defined by the retail chain. There is no variation in revenues and prices for stores from retail chain 4 and 5.

Table C 2. Average Store Prices of Food Categories (Dollars per Unit)

Item	Retail Chain 1		Retail Chain 2		Retail Chain 3		Retail Chain 4		Retail Chain 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Perishable Product										
Baked Goods	0.37	0.09	1.24	0.15	0.00	0.00	0.03	–	0.09	–
Beef	3.56	0.36	0.69	0.09	0.16	0.06	1.09	–	1.81	–
Beef/Veal	0.00	0.00							0.00	–
Chicken	1.49	0.20	0.19	0.03	0.02	0.01	0.46	–	0.93	–
Deli Cheese	0.22	0.05	0.22	0.04			0.06	–	0.14	–
Deli Meat	1.30	0.30	1.11	0.18			0.20	–	0.54	–
Deli Prepared	1.06	0.23	1.03	0.16	0.00	0.00	0.28	–	1.24	–
Fish	0.44	0.12	0.00	0.00			0.05	–	0.08	–
Fruit	7.39	0.71	7.22	0.61	3.85	1.20	3.65	–	5.89	–
Other Meat	0.38	0.07	0.21	0.03	0.09	0.03	0.16	–	0.39	–
Other Produce	0.31	0.05	0.21	0.02	0.05	0.03	0.11	–	0.19	–
Other Seafood	0.01	0.00	0.00				0.00	–	0.00	–
Pork	1.27	0.25	0.26	0.04	0.09	0.03	0.66	–	1.11	–
Sausage	0.00	0.00					0.01	–	0.01	–
Shellfish	0.17	0.06	0.07	0.05	0.00	0.00			0.00	–
Turkey	0.30	0.07	0.18	0.04	0.17	0.07	0.09	–	0.21	–
Vegetables	5.55	0.63	3.93	0.26	1.77	0.61	2.29	–	4.31	–
Processed Product										
Baby Food	0.72	0.67	1.85	0.26	3.81	1.35	1.59	–	1.43	
Bakery	5.10	0.36	3.37	0.21	3.53	0.30	3.79	–	3.15	–
Baking	2.98	0.24	2.86	0.08	2.95	0.22	3.47	–	3.49	–
Breakfast	2.27	0.12	2.92	0.11	3.44	0.33	4.19	–	3.29	–
Candy	1.94	0.23	4.75	0.59	9.09	1.78	8.95	–	5.49	–
Carbonated Soft Drinks	2.97	0.42	3.28	0.47	5.25	1.29	5.83	–	4.65	–
Coffee & Tea	2.11	0.20	3.34	0.19	5.01	0.66	3.67	–	3.01	–
Condiments & Sauces	3.25	0.16	2.97	0.09	2.75	0.19	3.47	–	3.59	–
Cookies & Crackers	2.57	0.24	3.18	0.26	4.00	0.51	3.80	–	2.89	–
Dairy	12.64	0.78	11.37	0.55	9.64	1.19	6.85	–	8.67	–
Drink Mixes	0.33	0.03	0.40	0.05	0.60	0.09	0.72	–	0.55	–
Ethnic	1.13	0.11	1.12	0.04	0.98	0.09	1.05	–	1.20	–
Frozen Baked Goods	0.37	0.05	0.17	0.03	0.16	0.04	0.30	–	0.33	–

Frozen Beverages	0.16	0.02	0.08	0.01	0.06	0.01	0.06	–	0.09	–
Frozen Desserts	2.22	0.16	1.70	0.12	1.94	0.34	1.98	–	1.61	–
Frozen Fruits & Vegetables	1.18	0.09	1.33	0.06	1.16	0.22	0.61	–	0.74	–
Frozen Meals	4.19	0.28	4.54	0.34	4.91	0.57	6.25	–	5.31	–
Frozen Meat/Poultry/Seafood	3.90	0.30	4.23	0.21	2.92	0.52	3.16	–	2.90	–
Frozen Snacks	0.34	0.06	0.41	0.05	0.47	0.08	0.75	–	0.70	–
Juices	1.06	0.20	1.18	0.06	1.55	0.18	2.32	–	1.57	–
Liquor	0.12	0.07	0.76	1.67	0.00	0.00	0.16	–	2.98	–
Meals	4.46	0.27	3.99	0.18	3.61	0.27	5.67	–	5.47	–
Non–Fruit Drinks	0.07	0.02	0.15	0.02	0.13	0.03	0.09	–	0.12	–
Other Frozen	0.00	0.00	0.01	0.00	0.02	0.01			0.00	–
Other Refrigerated	0.06	0.01	0.00	0.00	0.00	0.00	0.02	–	0.04	–
Produce	1.36	0.21	1.36	0.07	1.02	0.32	0.16	–	0.28	–
Refrigerated Baked Goods	0.36	0.05	0.17	0.03	0.08	0.02	0.10	–	0.12	–
Refrigerated Beverages	1.49	0.17	1.47	0.13	1.85	0.40	1.37	–	1.17	–
Refrigerated Condiments	0.80	0.15	0.85	0.08	0.51	0.15	0.22	–	0.37	–
Refrigerated Desserts	0.11	0.03	0.10	0.01	0.04	0.01	0.11	–	0.15	–
Refrigerated Dough	0.37	0.05	0.31	0.04	0.24	0.05	0.29	–	0.37	–
Refrigerated Meals	1.37	0.22	1.11	0.12	1.11	0.21	0.81	–	0.96	–
Refrigerated Meats	5.23	0.51	5.11	0.36	3.75	0.82	3.94	–	4.99	–
Snacks	5.11	0.37	8.59	0.47	11.70	1.41	9.58	–	7.27	–
Sports/Energy Drinks	0.71	0.10	0.78	0.08	1.30	0.19	2.35	–	1.57	–
SS Fruit	0.67	0.08	0.97	0.08	0.96	0.16	0.70	–	0.62	–
SS Vegetables	1.00	0.08	0.69	0.03	0.41	0.06	0.59	–	0.69	–
Water	1.45	0.19	2.00	0.15	2.99	0.41	1.93	–	1.20	–

Note: Retail chain 1, 2, and 3 report sales information at the store level, while retail chain 4 and 5 report at the Retail Market Area (RMA) level. The RMA is defined by the retail chain. There is no variation in sales revenues and prices for stores from retail chain 4 and 5.

Table C 3. Regression Results of Stores' Profit Margins on Number of Nearby Rivals within 8 km

Variable	Competition across Stores			Collusion within the Same Site			Collusion within the Retail Chain		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Constant	–28.014***	–28.381***	–26.425***	–44.360***	–45.042***	–40.597***	–17.796***	–16.813***	–20.882***
	(2.417)	(2.416)	(2.487)	(2.489)	(2.473)	(2.541)	(3.170)	(3.143)	(3.293)
Number of Nearby Stores	–0.013***			–0.013***			–0.010***		
	(0.001)			(0.001)			(0.002)		
Number of Nearby Stores		–0.020***			–0.025***			0.008**	
from Same Retail Chain		(0.003)			(0.003)			(0.004)	
Number of Nearby Stores		–0.010***			–0.008***			–0.017***	
from Other Retail Chains		(0.002)			(0.002)			(0.002)	
Number of Nearby Stores			–0.020***			–0.011***			–0.029***
from Retail Chain 1			(0.003)			(0.003)			(0.004)
Number of Nearby Stores			0.009*			0.003			0.021***
from Retail Chain 2			(0.004)			(0.005)			(0.006)
Number of Nearby Stores			–0.021***			–0.034***			0.008
from Retail Chain 3			(0.005)			(0.005)			(0.006)
Number of Nearby Stores			–0.013			–0.032***			–0.006
from Retail Chain 4			(0.008)			(0.008)			(0.011)
Number of Nearby Stores			0.011**			0.009*			0.008
from Retail Chain 5			(0.006)			(0.006)			(0.007)
Pop. Percentage of Marriage	–0.137***	–0.141***	–0.128***	–0.209***	–0.217***	–0.184***	–0.101***	–0.090***	–0.128***
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.020)	(0.020)	(0.021)
Pop. Percentage of Education	0.204***	0.211***	0.171***	0.369***	0.382***	0.303***	0.202***	0.183***	0.243***
Less than High Schl Degree	(0.028)	(0.028)	(0.030)	(0.028)	(0.028)	(0.030)	(0.036)	(0.036)	(0.039)
Pop. Percentage of Education	0.240***	0.228***	0.237***	0.274***	0.251***	0.255***	0.088**	0.121***	0.109***
High School Degree	(0.029)	(0.029)	(0.029)	(0.030)	(0.030)	(0.029)	(0.038)	(0.038)	(0.038)
Pop. Percentage of Education	0.155***	0.146***	0.142***	0.146***	0.128***	0.115***	0.051	0.077**	0.064*

Bachelor's Degree	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)	(0.026)	(0.034)	(0.034)	(0.034)
Pop. Percentage of Education	0.032***	0.031***	0.046***	0.098***	0.097***	0.121***	–0.012	–0.009	–0.017
Graduate Degree	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)	(0.015)	(0.016)
Pop. Percentage of HH Income	0.452***	0.468***	0.425***	0.757***	0.787***	0.697***	0.398***	0.354***	0.451***
≥ 25k and < 50k	(0.046)	(0.046)	(0.047)	(0.047)	(0.048)	(0.048)	(0.060)	(0.060)	(0.063)
Pop. Percentage of HH Income	–0.015	–0.009	–0.009	–0.004	0.007	0.012	–0.134***	–0.150***	–0.137***
≥ 50k and < 75k	(0.025)	(0.025)	(0.024)	(0.025)	(0.025)	(0.025)	(0.032)	(0.032)	(0.032)
Pop. Percentage of HH Income	0.350***	0.362***	0.320***	0.622***	0.644***	0.560***	0.319***	0.287***	0.360***
≥ 75k and < 100k	(0.039)	(0.040)	(0.040)	(0.041)	(0.041)	(0.041)	(0.052)	(0.051)	(0.053)
Pop. Percentage of HH Income	0.331***	0.345***	0.309***	0.558***	0.583***	0.501***	0.275***	0.238***	0.331***
≥ 100k	(0.038)	(0.038)	(0.039)	(0.039)	(0.039)	(0.040)	(0.049)	(0.049)	(0.052)
Retail Chain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272
Adjusted R^2	0.654	0.655	0.673	0.714	0.718	0.734	0.511	0.521	0.529
F Statistic	78.352	76.435	75.573	103.19	102.25	101.07	43.869	44.246	41.822

Note: The dependent variables are stores' profit margins computed under the alternative pricing scenario. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table C 4. Regression Results of Stores' Profit Margins on Number of Nearby Rivals within 10 km

Variable	Competition across Stores			Collusion within the Same Site			Collusion within the Retail Chain		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Constant	–24.439***	–24.442***	–20.190***	–40.879***	–41.335***	–36.206***	–15.808***	–12.451***	–12.087***
	(2.477)	(2.491)	(2.467)	(2.554)	(2.564)	(2.471)	(3.273)	(3.148)	(3.322)
Number of Nearby Stores	–0.011***			–0.011***			–0.007***		
	(0.001)			(0.001)			(0.001)		
Number of Nearby Stores		–0.011***			–0.015***			0.021***	
from Same Retail Chain		(0.002)			(0.002)			(0.003)	
Number of Nearby Stores		–0.011***			–0.009***			–0.020***	
from Other Retail Chains		(0.001)			(0.001)			(0.002)	
Number of Nearby Stores			–0.007***			–0.007***			–0.007**
from Retail Chain 1			(0.002)			(0.002)			(0.003)
Number of Nearby Stores			–0.001			–0.006			0.012**
from Retail Chain 2			(0.004)			(0.004)			(0.005)
Number of Nearby Stores			–0.033***			–0.035***			–0.023***
from Retail Chain 3			(0.004)			(0.004)			(0.005)
Number of Nearby Stores			–0.045***			–0.065***			–0.037***
from Retail Chain 4			(0.006)			(0.006)			(0.008)
Number of Nearby Stores			–0.005			0.001			–0.002
from Retail Chain 5			(0.004)			(0.004)			(0.006)
Pop. Percentage of Marriage	–0.126***	–0.126***	–0.096***	–0.198***	–0.201***	–0.167***	–0.097***	–0.074***	–0.070***
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.020)	(0.020)	(0.021)
Pop. Percentage of Education	0.142***	0.142***	0.072**	0.308***	0.318***	0.222***	0.164***	0.091**	0.116***
Less than High Schl Degree	(0.029)	(0.029)	(0.029)	(0.030)	(0.030)	(0.029)	(0.038)	(0.037)	(0.039)
Pop. Percentage of Education	0.207***	0.207***	0.195***	0.241***	0.235***	0.207***	0.059	0.106***	0.064*
High School Degree	(0.028)	(0.028)	(0.028)	(0.029)	(0.029)	(0.028)	(0.037)	(0.036)	(0.038)
Pop. Percentage of Education	0.124***	0.124***	0.095***	0.116***	0.112***	0.066***	0.023	0.052*	0.013

Bachelor's Degree	(0.025)	(0.025)	(0.025)	(0.026)	(0.026)	(0.025)	(0.033)	(0.032)	(0.033)
Pop. Percentage of Education	0.032***	0.032***	0.060***	0.099***	0.098***	0.131***	–0.009	–0.002	0.010
Graduate Degree	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)	(0.016)	(0.015)	(0.017)
Pop. Percentage of HH Income	0.417***	0.417***	0.340***	0.724***	0.736***	0.657***	0.387***	0.300***	0.303***
≥ 25k and < 50k	(0.047)	(0.047)	(0.047)	(0.048)	(0.048)	(0.047)	(0.061)	(0.059)	(0.064)
Pop. Percentage of HH Income	0.008	0.008	0.027	0.019	0.021	0.060***	–0.110***	–0.127***	–0.104***
≥ 50k and < 75k	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)	(0.023)	(0.031)	(0.029)	(0.031)
Pop. Percentage of HH Income	0.303***	0.303***	0.229***	0.576***	0.587***	0.499***	0.295***	0.215***	0.223***
≥ 75k and < 100k	(0.040)	(0.041)	(0.040)	(0.041)	(0.042)	(0.040)	(0.053)	(0.051)	(0.054)
Pop. Percentage of HH Income	0.294***	0.294***	0.222***	0.522***	0.533***	0.450***	0.261***	0.180***	0.192***
≥ 100k	(0.038)	(0.039)	(0.039)	(0.039)	(0.040)	(0.039)	(0.051)	(0.049)	(0.052)
Retail Chain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272
Adjusted R^2	0.658	0.658	0.679	0.717	0.717	0.749	0.510	0.551	0.522
F Statistic	79.813	77.256	77.770	104.71	101.72	109.29	43.712	49.806	40.731

Note: The dependent variables are stores' profit margins computed under the alternative pricing scenario. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.

Table C 5. Regression Results of Stores' Profit Margins on Number of Nearby Rivals within 16 km

Variable	Competition across Stores			Collusion within the Same Site			Collusion within the Retail Chain		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Constant	–26.576***	–26.830***	–30.865***	–42.530***	–42.831***	–47.729***	–18.221***	–19.707***	–20.931***
	(2.534)	(2.530)	(2.535)	(2.598)	(2.592)	(2.578)	(3.303)	(3.129)	(3.278)
Number of Nearby Stores	–0.005***			–0.006***			–0.002**		
	(0.001)			(0.001)			(0.001)		
Number of Nearby Stores		–0.002			–0.002			0.015***	
from Same Retail Chain		(0.001)			(0.001)			(0.002)	
Number of Nearby Stores		–0.007***			–0.007***			–0.010***	
from Other Retail Chains		(0.001)			(0.001)			(0.001)	
Number of Nearby Stores			0.011***			0.014***			0.012***
from Retail Chain 1			(0.002)			(0.002)			(0.003)
Number of Nearby Stores			–0.019***			–0.025***			–0.001
from Retail Chain 2			(0.004)			(0.004)			(0.005)
Number of Nearby Stores			–0.024***			–0.024***			–0.033***
from Retail Chain 3			(0.003)			(0.003)			(0.003)
Number of Nearby Stores			0.001			0.006			–0.026***
from Retail Chain 4			(0.006)			(0.006)			(0.008)
Number of Nearby Stores			–0.016***			–0.022***			–0.002
from Retail Chain 5			(0.004)			(0.004)			(0.005)
Pop. Percentage of Marriage	–0.134***	–0.133***	–0.154***	–0.202***	–0.201***	–0.229***	–0.111***	–0.107***	–0.122***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)	(0.021)	(0.020)	(0.021)
Pop. Percentage of Education	0.166***	0.164***	0.183***	0.327***	0.324***	0.361***	0.192***	0.178***	0.149***
Less than High Schl Degree	(0.029)	(0.029)	(0.029)	(0.030)	(0.030)	(0.030)	(0.038)	(0.036)	(0.038)
Pop. Percentage of Education	0.196***	0.204***	0.172***	0.233***	0.243***	0.202***	0.045	0.095***	–0.002
High School Degree	(0.029)	(0.029)	(0.032)	(0.029)	(0.030)	(0.032)	(0.037)	(0.036)	(0.041)
Pop. Percentage of Education	0.117***	0.122***	0.070**	0.113***	0.118***	0.060**	0.009	0.036	–0.072**

Bachelor's Degree	(0.026)	(0.026)	(0.028)	(0.027)	(0.027)	(0.029)	(0.034)	(0.032)	(0.037)
Pop. Percentage of Education	0.046***	0.051***	0.107***	0.111***	0.117***	0.175***	0.003	0.033**	0.084***
Graduate Degree	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)	(0.015)	(0.015)	(0.017)
Pop. Percentage of HH Income	0.449***	0.446***	0.530***	0.744***	0.740***	0.844***	0.434***	0.415***	0.504***
≥ 25k and < 50k	(0.048)	(0.048)	(0.048)	(0.050)	(0.049)	(0.049)	(0.063)	(0.060)	(0.062)
Pop. Percentage of HH Income	0.039*	0.042*	0.112***	0.045*	0.049**	0.121***	–0.081***	–0.062**	0.052
≥ 50k and < 75k	(0.023)	(0.023)	(0.026)	(0.024)	(0.024)	(0.026)	(0.030)	(0.029)	(0.034)
Pop. Percentage of HH Income	0.346***	0.345***	0.429***	0.611***	0.610***	0.718***	0.340***	0.333***	0.383***
≥ 75k and < 100k	(0.041)	(0.041)	(0.041)	(0.042)	(0.042)	(0.042)	(0.053)	(0.050)	(0.053)
Pop. Percentage of HH Income	0.319***	0.315***	0.380***	0.535***	0.531***	0.615***	0.302***	0.283***	0.341***
≥ 100k	(0.040)	(0.040)	(0.041)	(0.041)	(0.041)	(0.041)	(0.053)	(0.050)	(0.052)
Retail Chain Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272	1,272
Adjusted R^2	0.643	0.645	0.670	0.708	0.709	0.734	0.503	0.555	0.548
F Statistic	74.927	73.111	74.792	100.26	98.006	101.25	42.444	50.438	44.973

Note: The dependent variables are stores' profit margins computed under the alternative pricing scenario. Standard errors are presented in parentheses with asterisks indicating the level of significance, where *** represents the 1 percent level of significance, ** 5 percent, and * 10 percent.